THE PROTEIN RELATIONSHIPS IN RESPIRATION REGULATION ELUCIDATED BY X-RAY CRYSTALLOGRAPHY

by

Regina Dawn Kettering

Biochemistry and Molecular Biology, Lebanon Valley College, 2003

Submitted to the Graduate Faculty of

The Kenneth P. Deitrich School of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Regina Dawn Kettering

It was defended on

August 2, 2012

and approved by

Dr. Andrew P. VanDemark, Assistant Professor, Biological Sciences Department

Dr. Paula J. Grabowski, Professor, Biological Sciences Department

Dr. Anthony Schwacha, Associate Professor, Biological Sciences Department

Co-chair: Dr. Roger W. Hendrix, Professor, Biological Sciences Department

Co-chair: Dr. Joanne I. Yeh, Professor, Department of Structural Biology

THE PROTEIN RELATIONSHIPS IN RESPIRATION REGULATION ELUCIDATED BY X-RAY CRYSTALLOGRAPHY

Regina Dawn Kettering, PhD

University of Pittsburgh, 2012

Many pathways feed into the cellular respiration system. Hemoglobin supplies the oxygen necessary for aerobic respiration. Glycerol metabolism contributes electrons to the electron transport chain. Rap1B regulates superoxide production and platelet activation. I have determined 14 mutant structures that demonstrate how changes affect the regulation and function of the above systems: structural models of *Enterococcus casseliflavus* glycerol kinase mutants with glycerol or ethylene glycol, *Pseudomonas aeruginosa* glycerol kinase with glycerol or ethylene glycol, octameric human hemoglobin mutants, and two Rap1B mutants. The structures demonstrate that small changes in one region of the model can significantly affect the protein activity and adopted conformation of other portions of the model.

In the case of *En. casseliflavus* GK, the presence of loop interactions creates a domino effect that closes the active cleft and orients the substrates properly for glycerol phosphorylation. The His232Arg mutants with glycerol or ethylene glycol contain contacts between the activation loop and Domain I that stabilize a conformation that causes closure of the glycerol binding cleft. The *P. aeruginosa* GK adopts similar conformations to the *E. coli* structure in the glycerol-bound form. The activation cleft closes when bound to glycerol rather than ethylene glycol as in the *En. casseliflavus* structures, but requires ATP to stabilize the area close to the active site entrance.

The octameric structure of the hemoglobin mutants are affected by the presence of ligand in the molecule. The αAsn78Cys mutant does not show the same type of variability in the disulfide bridges, and therefore it does not have as much variability in the tetramers as the βGly83Cys mutants. The structures and crystallographic results suggest that the ligand stabilizes the βGly83Cys mutant relative tetramer motion.

The Rap1B models demonstrate changes in the switch II region determined by the phosphorylation of Ser179 rather than the presence of GTP. The negatively-charged Ser179Asp residue changes the interaction between the C-terminal loop and the core of the protein, presenting a binding area that is inaccessible in the Ser179Ala mutant.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my research advisor, Dr. Joanne I. Yeh, for her support and guidance during my graduate career. She has given me a series of interesting projects and has provided opportunities for me that would otherwise have been unattainable.

My committee members have been extremely supportive. I would like to thank Dr. Paula Grabowski, Dr. Roger Hendrix and Dr. Anthony Schwacha for their understanding and guidance. I would like to thank Dr. Andrew VanDemark for his willingness to step in on a very short notice.

Without the support of other researchers who helped me during graduate school, this research would not have been possible. Members of the Yeh lab assisted in my training both in research and in life. Dr. Ruth L. Saxl, Dr. Shoucheng Du, Dr. JohnJeff Alvarado, Dr. Haibin Shi, Dr. Unmesh Chinte, Dr. Laurie Betts and Dr. Ivan Vorontsov were always available and willing to answer questions. Dr. Chien Ho and his laboratory were very generous in providing me with octameric hemoglobin protein, and providing me with interesting feedback and discussions about the work. Dr. Daniel Altschuler and his laboratory provided Rap1B and interesting interpretations of the Rap1B models.

Last but not least I would like to express appreciation for my family and friends who were supportive of my dreams and encouraged me throughout my schooling. Special thanks

goes to Ms. Vanessa Kettering and Mr. John Orndorff, Jr, both of whom tolerated my laboratory schedule.

ABBREVIATIONS

Abbreviations that are used are shown with the full name and the abbreviation in parentheses for its first instance. All further instances use the listed abbreviation. Abbreviations that are found within this dissertation are as follows: 2,3-diphosphoglycerol (2,3-DPG); adenosine diphosphate (ADP); adenosine triphosphate (ATP); Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane (Bis-Tris); 1,3-Bis[tris(hydroxymethyl)methylamino]propane (BTP); 3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS); 2-(cyclohexylamino) ethanesulfonic acid (CHES); cyanide (CN); carbon monoxide (CO); carbon dioxide (CO_2); dihydroxyacetone phosphate (DHAP); *Enterococcus casseliflavus* (*En. casseliflavus*); *Escherichia coli* (*E. coli*); *Enterococcus faecalis* (*En. faecalis*); electron transport chain (ETC); flavin adenine dinucleotide (FAD); flavin adenine dinucleotide hydroquinone ($FADH_2$); glycerol-3-phosphate (G3P); guanosine diphosphate (GDP); glycerol kinase (GK); *Enterococcus casseliflavus* glycerol kinase complexed with *Enterococcus faecalis* histidine phosphorylating protein (GK:HPr); syn-glycerol-3-phosphate dehydrogenase (GlpD); glycerol facilitator (GlpF); syn-glycerol-3-phosphate oxidase (GlpO); glycerol-3- phosphate transporter (GlpT); guanosine triphosphate (GTP); adult human hemoglobin (HbA); hemoglobin-based oxygen carriers (HBOC); 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES); histidine phosphorylating protein (HPr); haptoglobin (Hpt); isopropyl β-D-1-thiogalactopyranoside (IPTG); Koshland, Nemethy, Filmer (KNF); monomethyl ether (MME); 2-(N-

Morpholino)ethanesulfonic acid (MES); Monod, Wyman, Changeux (MWC); nicotinamide adenine dinucleotide (NAD^+); nicotinamide adenine dinucleotide hydroquinone (NADH); nicotinamide adenine diphosphate hydroquinone (NADPH); oxygen molecule (O_2); octameric hemoglobin (oHb); *Pseudomonas aeruginosa* (*P. aeruginosa*); protein data bank (PDB); polyethylene glycol (PEG); phosphotransferase system (PTS); Root mean square deviation (RMSD); *Streptococcus pyogenes* (*S. pyogenes*); N-(Tris[hydroxymethyl]methyl)-3-aminopropanesulfonic acid (TAPS); tris(hydroxymethyl)aminomethane (Tris);

1.0 RESPIRATION BACKGROUND

Respiration is a broad term that applies to several biological processes. Cellular respiration is the catabolic process of breaking items such as glucose or glycerol down so that the cell can use the products to form energy. Respiration can also refer to the act of breathing and supplying oxygen to the various parts of the body. The electron transport chain (ETC) utilizes oxygen, carried by hemoglobin, as a final electron acceptor, splitting oxygen into two molecules of water. Oxygen from respiration can be shunted to create superoxide for destroying pathogens, which is controlled by Rap1 proteins. The electrons can be supplied by either nicotinamide adenine dinucleotide (NADH) or flavin adenine dinucleotide hydroquinone ($FADH_2$). Glycerol metabolism can also provide electrons for use in the ETC. This dissertation examines hemoglobin, the glycerol metabolism pathway, and superoxide production, all of which are involved in respiration.

1.1 HEMOGLOBIN AS A KEY MOLECULE IN RESPIRATION

Human hemoglobin has been studied for over a century. The individual components of the blood cells and the ability to form red crystals were first examined, followed by the structure and allosteric regulation of oxygen delivery [1]. Adult human hemoglobin consists of two $\alpha\beta$

dimers, creating a heterotetramer [2-7]. The α chains consist of 141 residues, whereas the β chains contain 146 residues. Although only 43% identical, the two chains fold into a nearly identical tertiary structure, and are also nearly identical in tertiary structure to myoglobin, despite only 26% and 25.5% amino acid identity between myoglobin and hemoglobin α and β, respectively [8]. Myoglobin serves as a monomeric oxygen carrier in muscle tissue [9]. Hemoglobin was the second crystal structure that was determined, revealing a quaternary structure that consisted of a dimer of interacting αβ dimers (see [3, 4] for interactions). The more stable dimerization interactions stabilize the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers, whereas interactions between dimers are listed as $\alpha_1\beta_2$ or $\alpha_2\beta_1$ interactions (see Figure 1.1).

Hemoglobin contains a prosthetic group heme (see Figure 1.1). Heme contains four imidazole rings linked together with the nitrogen atoms pointing towards the center of the ring. An iron atom is coordinated by the nitrogen atoms. Portions of the carbon ring contain methyl or ethyl groups that further enhance the hydrophobicity of the heme group. Two carboxylate groups are attached to two of the four imidazole groups, oriented on the same side of the porphyrin ring. The carboxylate groups create a hydrophilic area that helps the heme orient in the hydrophobic pocket of the α and β monomers of hemoglobin. Each monomer contains one heme group, which stabilizes the hemoglobin tertiary structure [2-4, 6, 7, 10-12]. During oxygen transport, one molecule of oxygen binds to each heme group. Without the heme group, hemoglobin would be unable to carry oxygen. Carbon monoxide, another two-atom molecule with similar electron properties to oxygen, can also bind to the heme groups [13, 14]. Carbon monoxide binding is much tighter than oxygen binding in either free or bound heme, but the presence of protein in bound heme significantly decreases the relative affinity of carbon monoxide [15]. The iron present in heme only binds to oxygen or carbon monoxide when in a

reduced +2 state. When oxidized, heme does not bind to either oxygen or carbon monoxide, but it does bind to cyanide.

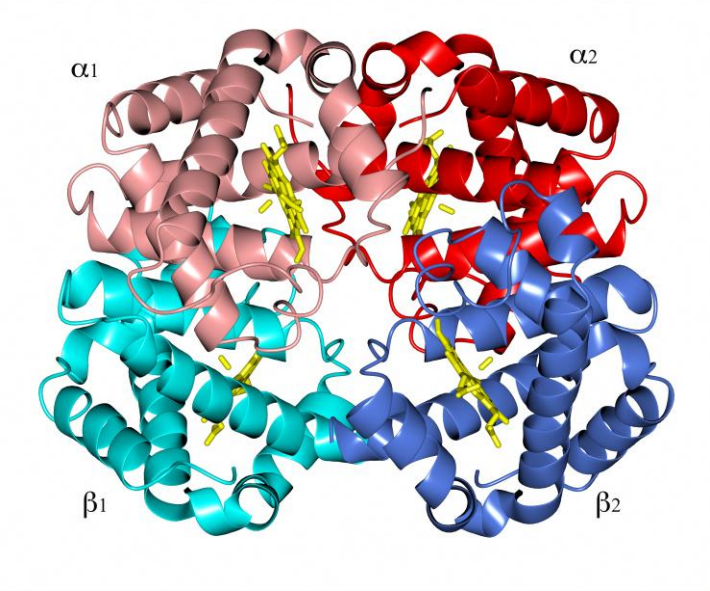

Figure 1.1 Tetrameric HbA contains 2 α subunits (in shades of red) and 2 β subunits (in shades of blue). Each monomer contains a heme group (in yellow). A yellow carbon monoxide molecule is shown interacting with each heme group.

Intimate interactions between the α_1-β_2 monomers strongly affect the uptake and release of oxygen. Initial models of allosteric regulation proposed, in conjunction with Perutz's hemoglobin models, that salt bridges and hydrogen bonds between the α_1-β_2 monomers restrained the quaternary structure into the T conformation [2-4]. The T conformation is considered to be the Tense state, as the lack of a sixth ligand (normally formed by the oxygen molecule) is absent and allows the heme iron to deform the heme porphyrin plane [2]. The R state is considered to be the Relaxed state because the iron is in the same plane as the porphyrin carbon and nitrogen atoms [2]. Adult human hemoglobin (HbA) adopts the R state when liganded with oxygen, carbon monoxide, cyanide or water when oxidized [2-4, 6, 7, 12, 16, 17]. Additionally, mutations along the interface can drastically affect the cooperativity and oxygen

affinity. Mutations are listed as being the from the α or β monomer with the mutant created listed as for other proteins. For example, the mutation βTrp37Ala (a mutation on the β chain at residue 37 that replaces Trp with Ala) removes critical interactions in the $\alpha_1\beta_2$ interface, increasing the oxygen affinity and drastically reducing the cooperativity [6]. However, mutations that decrease the oxygen affinity correlate to an increased heme oxidation [18].

1.1.1 Modulating Oxygen Affinity to Create a Physiologically Relevant Oxygen Carrier

In order to decrease the oxygen affinity of HbA in red blood cells, 2,3-diphosphoglycerol (2,3-DPG) is created by the body. 2,3-DPG binds in the cavity between the β subunits created in the ligand-free T-state of HbA. The area of this cavity is surrounded by positively-charged amino acids βLys82, βHis2, βHis143 and the NH_3^+ terminus of βVal1 [11, 19, 20]. The effects exerted by these small molecules finely tune the oxygen affinity of HbA, so that the delivery of oxygen is tightly optimized to physiological partial pressures, osmotic, and pH conditions [21, 22]. Specific changes in the configuration of residues at the porphyrin binding site when effectors are bound further augment the oligomeric interactions that lead to cooperativity, resulting in the modulation of oxygen (O_2), carbon monoxide (CO), and nitric oxide (NO) binding affinities [2, 10, 23-25]

During oxygen binding, HbA undergoes significant conformational shifts [3, 4, 6, 7]. The unliganded form is considered to be in the T state, or tense state, as the heme group is in a conformation that pushes the iron atom out of the plane of the porphyrin group [2]. The liganded form consists of the heme iron having six interactions, rather than the five seen in the T state [2]. The interactions including the new interaction with the ligand in the heme pocket, bring the iron into the plane of the porphyrin ring, resulting in the R state, or relaxed state [2]. Demonstrated

interactions in the x-ray crystal T state are broken during the T to R transition (see [3, 4, 6, 7] for details). The hydrogen bonds are proposed to stabilize the T state conformation.

1.1.2 Models of allostery

Two main models of hemoglobin allostery were introduced during the mid 1960's: the Koshland, Nemethy, Filmer model (KNF) known as the sequential model [26], and the Monod, Wyman, Changeux model (MWC) known as the concerted model [27]. Both models are based on the presence of two affinity states of hemoglobin, a low-affinity T state usually seen in the unliganded hemoglobin, and a high-affinity R state usually seen in the liganded hemoglobin. In the sequential model, the binding of one ligand to one monomer creates a structural change that is communicated to the surrounding monomers and increases their affinity for binding the next successive ligand. The tetramer can contain a combination of monomers in the T state or the R state. In the concerted model, binding of ligands puts energetic stress on the T state, causing the entire tetramer to become R state in the process of binding successive ligands. The concerted model is supported in part by crystallographic data, as no x-ray crystal structure to date simultaneously contains the structural T state and R state. Current models have taken the concerted model, and proposed that the structural and energetic transition between the T and R states are modulated by the binding of one oxygen molecule to one $\alpha\beta$ dimer, and a second oxygen molecule to the other $\alpha\beta$ dimer [28, 29]. Yonetani and coworkers are also supporting a model of allostery that suggests a separation between the structural state of T and R and the affinity state of T and R [16, 30].

A series of models have been proposed since the MWC and KNF models, some of which were detailed by Eaton, et al. [31]. The symmetry rule model proposed by Ackers and

coworkers is a combination of the two models, whereby the ligand state of one monomer affects the affinity of the monomer close to it to bind oxygen, and one ligand in each αβ dimer must be bound in order to switch from the T to the R state [32]. All of these models seek to accommodate the differences in the MWC 2-state model with the reality of ligand binding to hemoglobin. Many of the models create a situation where the last ligand to bind to hemoglobin has unreasonable affinity (sometimes a negative affinity). Current understanding of the biochemical models is necessary in order to create an appropriate blood substitute.

1.1.3 Blood substitutes

As the major oxygen carrier in mammalian systems, decreases in blood hemoglobin levels will affect the ability to carry out respiration. Current blood replacements consist of blood transfusions, which have storage requirements that limit their use in military and emergency situations [33]. Our collaborators at Carnegie Mellon University have created a series of octameric hemoglobins to be used to create a cell-free hemoglobin based oxygen carrier (HBOC) [34]. Here we discuss the effects these mutations have on the structure and biochemistry of hemoglobin.

1.2 GLYCEROL METABOLISM FEEDS INTO RESPIRATION

Glucose is a common "starter molecule" for respiration. The 6-carbon glucose gets broken down in aerobic respiration to smaller molecules, eventually becoming carbon dioxide and donating

electrons to the electron transport chain (ETC) via nicotinamide adenine dinucleotide (NAD^+) or flavin adenine dinucleotide (FAD) in the Kreb's cycle [35]. Electrons from citric acid, α-ketogluterate and malate reduce NAD^+ to nicotinamide adenine dinucleotide hydroquinone (NADH). The electrons from succinate are transferred to FAD to form flavin adenine dinucleotide hydroquinone ($FADH_2$). NADH and $FADH_2$ then donate their electrons to Complex I and Complex II of the ETC, respectively. The electrons travel through the ETC, pumping protons to create a proton gradient, and finally end up reducing O_2 and splitting it to create two water molecules.

In the absence of glucose, many organisms can use other carbon sources, such as glycerol. Glycerol is a 3-carbon molecule and can enter the aerobic respiration pathway as pyruvate. Glycerol can also be used through the glycerol metabolism pathway. Bacteria often utilize the glycerol metabolism pathway in the absence of phosphotransferase system (PTS) sugars such as glucose and fructose [36, 37]. Some bacterial strains do not grow well or do not grow at all if glycerol is not present [36, 38-41].

1.2.1 Glyerol Metabolism Pathway

The glycerol metabolism pathway starts with glycerol entering the cell via facilitated diffusion using the GlpF transporter. Once glycerol is inside the cell, it is phosphorylated by glycerol kinase (GK), creating a negatively charged glycerol-3-phosphate (G3P) that stays in the cytoplasm. G3P can also enter the cell via the GlpT transporter. G3P is dehydrogenated by GlpD or GlpO, depending on the bacteria. GlpD is a periplasmic membrane protein the uses FAD to pull the electrons from G3P and create dihydroxyacetone phosphate (DHAP) [42, 43].

GlpO performs a functionally similar process, but is a cytosolic protein [44]. The resulting electron transfer can be used to initiate the electron transport chain as in the case of glucose.

1.2.2 Regulation of Glycerol Metabolism

Bacteria can use multiple carbon sources during growth, but have preferential pathways. The most commonly-used pathway is the phosphoenolpyruvate pathway, in which glucose is brought into the cell, phosphorylated, and broken down into energy and other organic compounds [36]. These are listed at the PTS pathway, and are regulated via a series of enzyme phosphorylations, resulting in glucose gene expression [45]. However, when glucose is not present, other carbon sources, such as glycerol, can be used for energy and creating organic compounds [46].

The glycerol pathway is important for regulating growth in bacteria. When glycerol-3-phosphate (G3P) is present, a repressor is turned off, allowing expression of the glp operon and genes that allow bacteria to trap and utilize glycerol. The glp operon consists of the *glpK*, *glpF*, *glpD, glpR, glpT* and *glpQ [47-50]*. *GlpF* encodes a glycerol facilitator GlpF that allows glycerol to enter the cell via facilitated diffusion [51]. Without the gene product from *glpK*, glycerol kinase, the glycerol would leave the cell. Glycerol kinase (GK) phosphorylates glycerol to G3P, trapping it inside of the cell and lowering the glycerol concentration inside the cell so that more glycerol can enter. G3P created by GK gets shuttled to sn-glycerol-3-phosphate dehydrogenase (GlpD), the product of *glpD*, where it is converted to dihydroxyacetone phosphate (DHAP). DHAP can be used as an energy source or shuttled through pathways to create complex organic compounds necessary for cell growth. In some cases, the glycerol pathway is necessary for cell growth, and mutants that do not contain the glp operon are inviable

[37, 46, 49]. Although the pathway is found in both Gram positive and Gram negative bacteria, they are controlled very differently between the two groups.

1.2.3 Control of the Glp pathway in Gram negative bacteria

In Gram negative bacteria, GK activity is downregulated when fructose 1,6-bisphosphate (FBP) binds to loops present at the top of Domain I [52, 53] (see Figure 1.2 for domain structure). The interaction stabilizes an inactive tetramer [53]. This is different from the interactions with unphosphorylated enzyme IIA for glucose ($EIIA^{Glc}$), which also decreases activity of GK [54]. The unphosphorylated $EIIA^{Glc}$ interacts with portions of GK ~30Å away from the active site at residues 474-479 [54-56]. Phosphorylated $EIIA^{Glc}$ will not interact with GK because phosphorylation site would be buried in a hydrophobic pocket without enough room to account for the highly charged moiety [56].

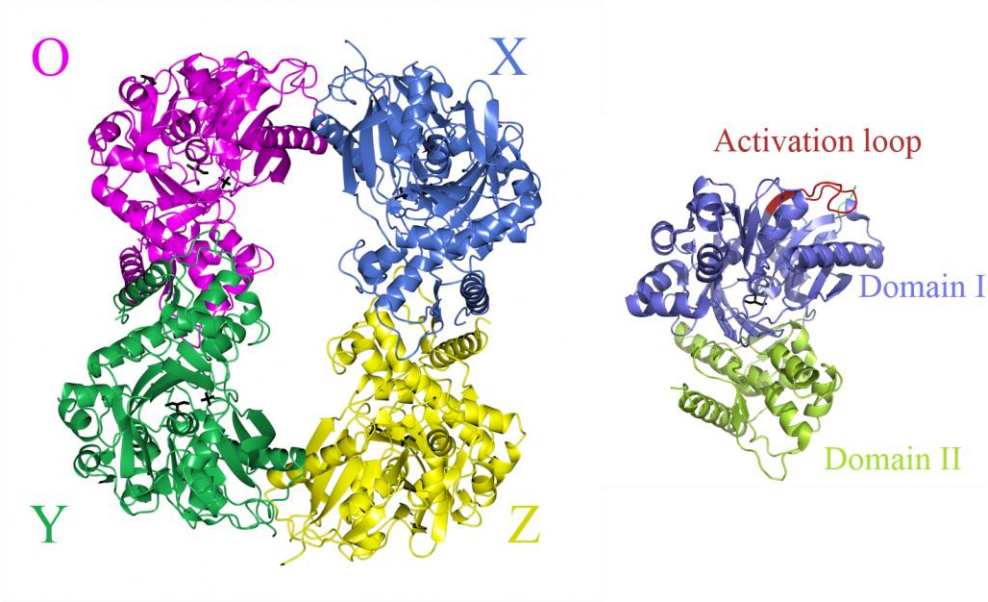

Figure 1.2 The tetrameric nomenclature and positions of GK (left). The image on the right shows Domain I (green), Domain II (blue) and the activation loop where phosphorylation occurs in Gram positive bacteria (red).

Glycerol kinase is in a structural family with hexokinase, actin, and hsp70. These structures contain an ATPase fold and a binding pocket for the substrate [54]. The GK models consist of two domains, I and II, which can be further split into three subsections. In GK the subsections IA, IB, IIA and IIB are found in the other members of its family. However, the subsections IC and IIC are only found in GK. These subsections are necessary for interaction with HPr and $EIIA^{Glc}$ [54].

1.2.4 Control of the Glp pathway in Gram positive bacteria

In the Gram positive bacteria, the activity of GK is modulated by directly phosphorylating a histidine on the activation loop. The presence of G3P and lack of PTS sugars initiate a phosphorylation cascade that phosphorylates a histidine phosphorylating protein (HPr) and allows it to pick up an extra phosphate to phosphorylate the activation loop of GK in Gram

positive bacteria [46]. The activation loop corresponds to the residues that are used for FBP binding in Gram negative bacteria. The discrepancy is caused by the fact that the Gram negative and Gram positive bacteria use a different orientation of dimers to regulate GK activity. In *E. coli*, the OY dimer is the physiologically relevant dimer, and stabilization of the tetramer prevents glycerol phosphorylation [57, 58]. Disruption of the OX dimer in *E. coli* GK by addition of large bulky residues in the OX dimer interface reduce or eliminate inhibition by FBP but do not inhibit activation [57, 58]. Previous attempts at crystallizing the Gram positive *Enterococcus faecalis* GK had been unsuccessful [59]. The Deutscher lab had created a series of mutants with the desire to examine how the particular mutations affected GK activity [59]. The Deutscher lab mutated *E. casseliflavus* GK to modify the phosphorylated histidine, His232. The His232Ala mutant has the same activity level as the basal activity level of the wild type (WT) GK. The His232Glu mutant adds a negative charge to mimic the charge of the phosphorylated histidine [59]. However, the activity level decreases -2.5-fold rather than increasing as expected [59]. A His232Arg mutant was also created, with a positive charge on the position of the regulating histidine, but this mutant has an increased activity. The 12-fold increase in activity is more than the 8-10-fold increase in activity seen in the activated phosphorylated GK [59]. The structures and the mutations' affects on the activity are discussed in this dissertation.

1.3 RAP1B REGULATION AND OXIDATIVE BURST RESPONSE

1.3.1 G-protein function

G-proteins are a family of small proteins between 16-21 kDa in size that are regulated by the presence of guanosine trophosphate (GTP) or guanosine diphosphate (GDP). When bound to GTP, the proteins are activated and bind to a number of cellular response proteins. G-proteins contain two switch regions, switch I and switch II, that change in response to GTP hydrolysis and regulate binding with effector proteins [60-62].

1.3.2 Ras superfamily of G proteins

The Ras superfamily contains proteins from the Ras, Rab, Rho, Arf, Rap, Ran, Rheb and Ral families, all of which are G proteins. These proteins have approximately 50% sequence homology to each other, and contain a high amount of structural similarity. Ras activates the MAPK pathway (reviewed in [63, 64]). Many proteins in the Ras superfamily are involved in regulation of cellular processes, including intracellular trafficking, actin cytoskeletal remodeling, cell polarity and cell cycle progression [65-67]. Ras is involved in the respiration pathway. The Gly12Val Ras mutant has been shown to retain GTP in the active site rather than hydrolyzing it [68]. The Gly12Val mutant switches the mitochondrial system from oxidative phosphorylation to glycolysis [69].

In the superoxide response, molecular oxygen is reduced by the nicotinamide adenine diphosphate hydroquinone (NADPH) oxidase complex to form superoxide ($O_2^{-}\bullet$) [70, 71]. Rap1A, a G-protein very similar to Ras, has been shown to affect the production of superoxide in

human B lymphocytes [72, 73]. Both Rap1A and Rap1B have been shown to be phosphorylated at this C-terminus at Ser residues by protein kinase A [73-79]. Both Rap1A and Rap1B contain hydrophibic post-translational modifications that cause them to be associated with the cell membrane [80, 81]. Rap1B in particular is isoprenylated at Cys181 [82]. However, phosphorylation of Rap1B Ser179 causes the protein to dissociate from the membrane and become cytosolic [83, 84]. The dissociation is proposed to affect activation and recruitment of secondary proteins to the membrane [82]. Rap1B activation has also been shown to induce platelet activation (reviewed in [85]).

Many protein data bank (PDB) models have demonstrated the Ras protein core that consists of a mixture of alpha and beta structures. The beta sheets have been shown to interact with various regulators [86-89]. The models show two switch regions that are involved in protein-protein interactions and regulation. The G-proteins also contain a GTP nucleotide binding site and a divalent cation binding site (usually magnesium).

Rap1B is extremely similar to Rap1A, with only the C-terminal tail comprising the difference. However, all structures published thus far contain a terminated Rap1 protein, with a deletion between residues 166-168 [87-89]. Studies of Ras demonstrate long-distance interactions are necessary for dissociation of GDP from the active site [90]. Two mutants have been created to elucidate the means by which the phosphorylation affects the structure, and therefore the activity of Rap1B. These are the first structures of the full-length Rap1B protein.

2.0 CONTROL OF GLYCEROL METABOLISM

2.1 INTRODUCTION

People with cystic fibrosis suffer from a buildup of liquid in their lungs, resulting in a propensity for infections. The excess liquid makes it difficult to clear the lungs of extraneous foreign materials, including bacteria. After contracting pneumonia, infections with bacteria such as *Pseudomonas aeruginosa* (*P. aeruginosa*) can become problematic. A challenge of treating *Pseudomonas* infections in lungs is their tendency to form biofilms, bacterial colonies that are connected by a film of polysaccharides, rhamnose, etc. [91]. Glycerol can be used to create multiple sugars so preventing glycerol kinase activity may also be used to prevent biofilm formation [92]. Syn-glycerol-3-phosphate dehydrogenase (GlpD), one of the proteins in the glycerol metabolism pathway, has been shown to affect biofilm production and persister cells [49]. While activation has been examined in Gram negative *E. coli*, this mechanism has not been previously determined for Gram positive pathogens such as *Enterococcus faecalis* (*En. faecalis*). Glycerol kinase (GK) from both Gram positive and Gram negative bacteria has a distance of approximately 25-30 Å from the activation or repression site to the glycerol binding site (see Figure 1.2). In Gram negative systems, $EIIA^{Glc}$ interacts with GK to prevent activation by binding to Domain II near the OY dimer interface (see Figure 1.2) [54]. When $EIIA^{Glc}$ is phosphorylated, it dissociates from GK, activating the protein [93]. In Gram positive systems, a

histidine on an activation or regulation loop is phosphorylated by a histidine-phosphorylating protein (HPr), activating GK [94]. The glycerol binding site and ATP binding site are situated between Domain I and Domain II [54]. Glycerol binds further into the activation cleft and must bind to GK first, after which ATP can bind close to the activation cleft entrance for phosphoryl transfer [54]. The phosphorylated histidine stays phosphorylated throughout the process, and can keep GK active until the phosphate dissociates from the histidine [94]. Once GK phosphorylates glycerol to glycerol-3-phosphate (G3P), it transfers G3P to either GlpD or GlpO [95, 96]. In order to understand this regulation process and the differences in the control mechanisms, examination of the Gram positive structures and the interactions between GK, HPr, GlpD and GlpO are necessary.

2.1.1 Mutants that affect GK activity

The Deutscher lab created a variety of mutations that altered the activity of the Gram positive *Enterococcus casseliflavus* (*En. casseliflavus*) GK. Previous mutations in the *Bacillus subtillus* GK altered the activity contradictory to the expectations [97]. The Deutscher lab mutated *E. casseliflavus* GK to modify the phosphorylated histidine, His232. The His232Ala mutant has the same activity level as the basal activity level of the wild type (WT) GK. The His232Glu mutant adds a negative charge to mimic the charge of the phosphorylated histidine [59]. However, the activity level decreases -2.5-fold rather than increasing as expected [59]. A His232Arg mutant was also created, with a positive charge on the position of the regulating histidine, but this mutant has an increased activity. The 12-fold increase in activity is more than the 8-10-fold increase in activity seen in the activated phosphorylated GK [59]. We crystallized and determined the structures of the *En. casseliflavus* His232Arg and His232Glu mutants in the

presence of glycerol or ethylene glycol in order to determine how these mutations affect the structure and activate *En. casseliflavus* GK. In order to complete a secondary comparison, we also crystallized and determined the structures of *P. aeruginosa* GK in the presence of glycerol or ethylene glycol. The glycerol-bound state shows the substrate-protein interactions that are necessary for activation, while the ethylene glycol-bound state shows an intermediate binding state of GK between the glycerol-bound state and the unbound state. These are the first structures that give insight into GK activation in Gram positive bacteria, and demonstrate the processes that occur during glycerol binding.

2.1.2 GK Regulation and Protein-Protein Interactions

In order to determine the structural components of the GK-HPr interaction, I attempted to cocrystallize the *En. casseliflavus* GK with the *En. faecalis* HPr. The *En. faecalis* and *En. casseliflavus* HPr are 88% identical, with 97% homology [98]. The resultant structures give insight into the regulation of the *En. casseliflavus* glycerol pathway.

I also examined the interaction between *P. aeruginosa* GK and GlpD or GlpO variants from different bacterial species. The Yeh lab was the first laboratory to determine the crystal structure of the *E. coli* GlpD, a monotopic membrane protein [99]. Both GlpD and GlpO pull electrons from the substrate G3P and transfer them to either NAD^+ or FAD. Electrons are then transferred to ubiquinones, which are associated with GlpD.

2.2 MATERIALS AND METHODS

2.2.1 Cloning of *Enterococcus casseliflavus* glycerol kinase

Dr. Ruth L. Saxl cloned the En. casseliflavus GK from the plasmids generously provided by the Deutscher lab. PCR used 200 ng of DNA template, 1.5 mM MgSO4, 0.2 mM dNTPs, 1X KOD buffer, KOD polymerase 1U/µL, 0.3 µM 3' primer, 0.3 µM 5' primer. After a 5 min 95° start, 20 cycles of 95° 1 min, 1.5 min 43°, 2 min 70°, followed by 20 cycles of 1 min 95°, 0.5 min 69°, 2 min 70° were followed by 10 min 70° and 4° cool. The product was blunt ligated into TOPO blunt vector using 1 µL TOPO-blunt, 4 µL pCR product, 1X T4 buffer and 1 µL T4 ligase in a total of 10 µL. Vector was used to transform chemically competent DH5α cells.

2.2.2 Cloning of *Pseudomonas aeruginosa* glycerol kinase

The *P. aeruginosa* GK gene was cloned by Dr. Ruth L. Saxl using genomic *P. aeruginosa* DNA, ATCC 15442, and primers of 5' CC CAT ATG CAT CAC CAT CAC CAT CAC CAT CAC GAA ACC CTG TAT TTT CAG GGC AAC GAC AAG CAC AAC AAG AAA T 3' that includes a 8-His N terminal tag, a TEV site, and an NdeI cleavage site, and a second primer of 5' GA CTC GAG TCA CAG TTC GCC GTC GTC C 3' that includes the stop codon and XhoI cleavage site for insertion into the plasmid. A second set of primers were needed for the middle of the *P. aeruginosa* GK, as the gene contained an internal XhoI site that would render only a partial GK gene if not mutated. The resulting mutation, a Glu209Glu mutation, was introduced

via a primer with the sequence 5' GGA TGC TCG AAG TGC TCG ACA 3' and complementary primer. The cloned segment was placed in a TOPO vector, a blunt-end ligation, and TOP10 cells were transformed to allow for propagation of the vector. The vector was sequenced to determine that the sequence is correct, and then was purified via a Qiagen Miniprep kit. The GK insert was digested out of the TOPO vector via BamHI and HinDIII endonuclease sites and ligated into pQE30-Xa vector between the BamHI and HinDIII restriction sites. pQE30-Xa containing *P. aeruginosa* GK was sequenced and used to transform XL1Blue cells for expression and purification.

2.2.3 Protein expression

GK was expressed by Dr. Ruth L. Saxl and me as in Yeh, et al. and Yeh, et al. [100, 101]. Overnight cultures of BL21DE3 pLysS were grown in Luria broth (LB, 10 g sodium chloride, 10 g tryptone, 5 g yeast extract per liter) and 50 µg/mL carbenicillin, 50 µg/mL chloramphenicol. The overnight culture was centrifuged at 4000 rpm in eppendorf 5910R centrifuge for 10-30 min, the LB decanted, and resuspended in enough LB plus carbenicillin and chloramphenicol to inoculate LB with a 1:100 ratio. Bacteria were grown to a density of 0.6-1.0 OD_{600},after which protein expression was induced for 4 hr using a final concentration of 1 mM IPTG. Cells were centrifuged for 20-30 min using a Sorvall RC3C Plus centrifuge at 4000 rpm. Cells were resuspended in 50 mM Tris pH 7.4, and then centrifuged again for 10-30 min at 4000 rpm using the ependorf 5810R centrifuge. Cells were frozen at -20°C or -80ºC until use.

2.2.4 Purification of His-tagged Protein

Cells were thawed on ice, suspended in 20-50 mL of 50 mM Tris pH 7.4-7.5 with 1X His-tagged protease inhibitor from Sigma-Aldrich and broken open using three passes through a fluidizer. The buffer for the fluidizer was 50 mM Tris pH 7.4. The suspension of broken cells was centrifuged for 30 min at 4000 rpm in ependorf 5810R centrifuge, and then further centrifuged for 1 hr at 100,000 G in a Sorval WX Ultra with a Ti70 rotor. This eliminated the cell debris and membrane components. The supernatant was put onto a cobalt column either by incubating with cobalt beads on a nutator to keep the 1:1 ratio of beads:liquid suspended overnight at 4°C or by flowing them over the beads using a gravity flow column. The supernatant was loaded onto a either a nickel column or a cobalt column, washed with wash buffer and eluted with elution buffer. Wash buffer for the nickel column consisted of 50 mM Tris pH 7.5, 300 mM sodium chloride, and 20 mM imidazole. Elution buffer for the nickel column consisted of 50 mM Tris pH 7.5, 300 mM sodium chloride, and 300 mM imidazole. Wash buffer for the cobalt column contained 50 mM Tris pH 7.5, 150 mM sodium chloride and 10 mM imidazole. Elution buffer for the cobalt column contained 50 mM Tris pH 7.5, 150 mM sodium chloride, 150 mM imidazole. The column was washed with 10 times the column volume in wash buffer. Protein was eluted using four times the column volume of elution buffer. The protein was diluted at least 10X with 50 mM Tris pH 7.5, then buffer exchanged and concentrated to 8-20 mg/mL before being flash frozen and stored at -80°C until use. Purification was preferentially done using a cobalt column that requires lower imidazole concentrations to elute the protein because the GK is unstable in high concentrations of imidazole and has a tendency to precipitate. GK was concentrated to 8-20 mg/mL and buffer exchanged into 50 mM Tris pH 8.0 using Amicon concentrators with 30 kDa cutoff. HPr was concentrated to 8-20 mg/mL and buffer exchanged

into 20 mM Tris pH 8.0 using Amicon concentrators with 3-5 kDa cutoff. Protein was separated into 50-200 µL aliquots, flash frozen in liquid nitrogen and stored at -80°C.

SDS-PAGE gels were run at each step to follow the protein purification steps. Previous work by Dr. Ruth L. Saxl and Dr. Shoucheng Du had shown that the *En. casseliflavus* GK appears on an SDS-PAGE gel at 55 kDa, and E. faecalis HPr runs on an SDS-PAGE gel around 9 kDa. 1X loading buffer of 2 mM BME, 2% SDS, 5% glycerol, 50 mM Tris-HCL pH 6.8 and bromphenol blue was added to 15 µL of sample. The sample was incubated at 90°C for 15 min, then loaded into wells of a discontinuous SDS-PAGE gel, with a stacking gel of 4% polyacrylamide/bis-acrylamide on top of the separating gel. Gels were run at constant voltage of 120 V for approximately 1-1.5 hr. Gels were stained with Coomassie blue stain and destained with 10% ethanol, 40% acetic acid. Gels showing the purified *En. casseliflavus* GK and *En. faecalis* HPr are in Figure 2.1 and Figure 2.2.

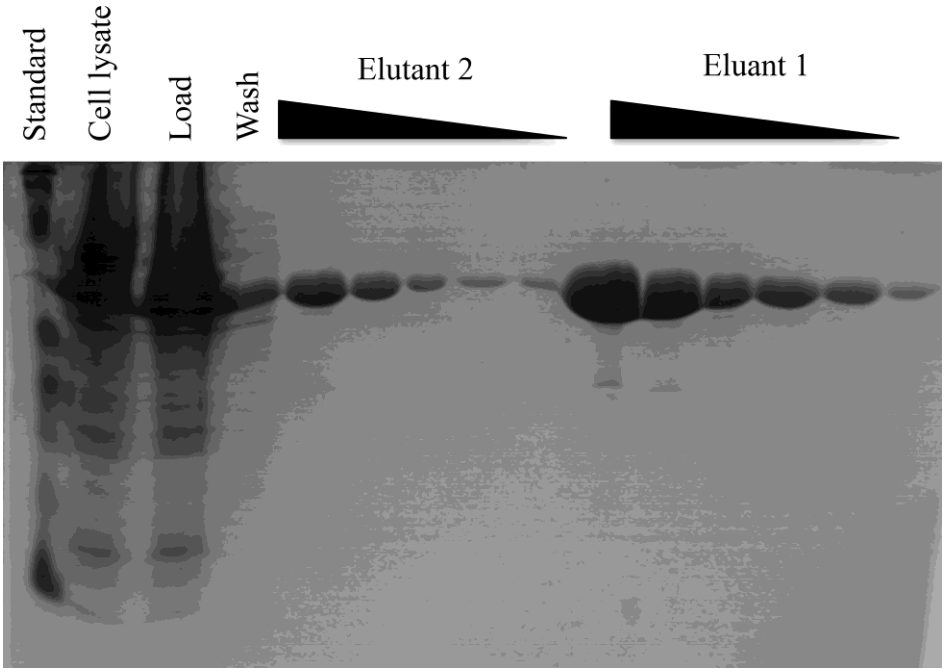

Figure 2.1 15% SDS-PAGE gel of his-tagged *En. casseliflavus* GK purification.

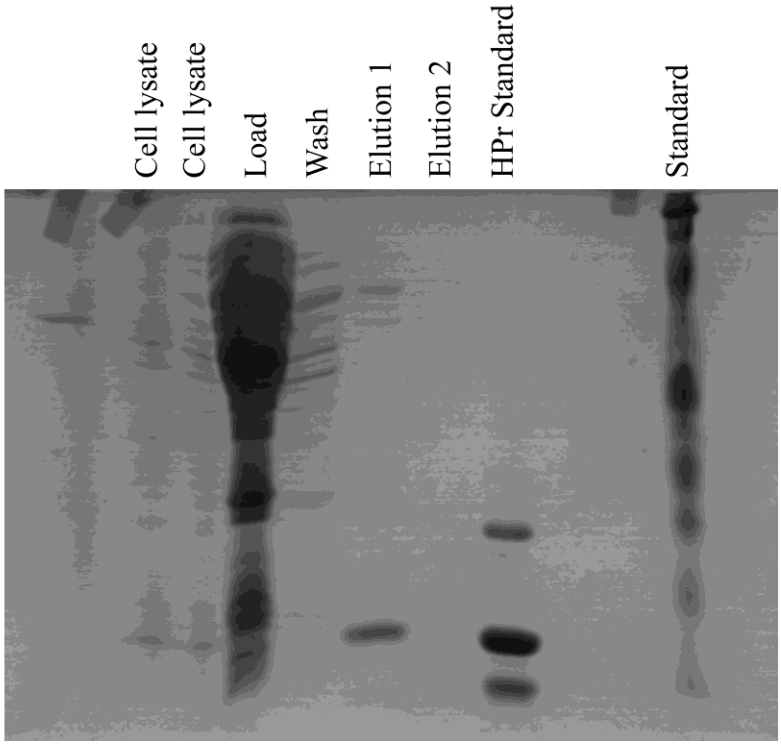

Figure 2.2 15% SDS-PAGE gel of his-tagged *E. faecalis* HPr purification. Elution 1 shows purified HPr, which was used for crystallization and SPR trials.

2.2.5 Non-his-tagged Purification

En. casseliflavus GK was purified as detailed in Yeh et al. and Yeh et al. [100, 101]. Cells were treated the same as those containing the plasmid with a His-tagged GK (see Section 2.2.4) through centrifugation of the cell debris at 100,000 G and collection of the supernatant. Supernatant was loaded onto DEAE columns and eluted using a 0 M to 1 M sodium chloride gradient in 20 mM Tris pH 8.0. Fractions containing GK were pooled, dialyzed into 20 mM Tris pH 8.0 to remove the salt and loaded onto a MonoQ column. Protein was eluted from the MonoQ column using a gradient of 0 M to 1 M sodium chloride (see Figure 2.3). Fractions containing purified GK were pooled, concentrated and buffer exchanged into 20 mM Tris pH 8.0

in an Amicon spin concentrator with a molecular weight cutoff of 30 kDa. Protein was flash frozen in aliquots of 50-200 µL and stored at -80°C until use.

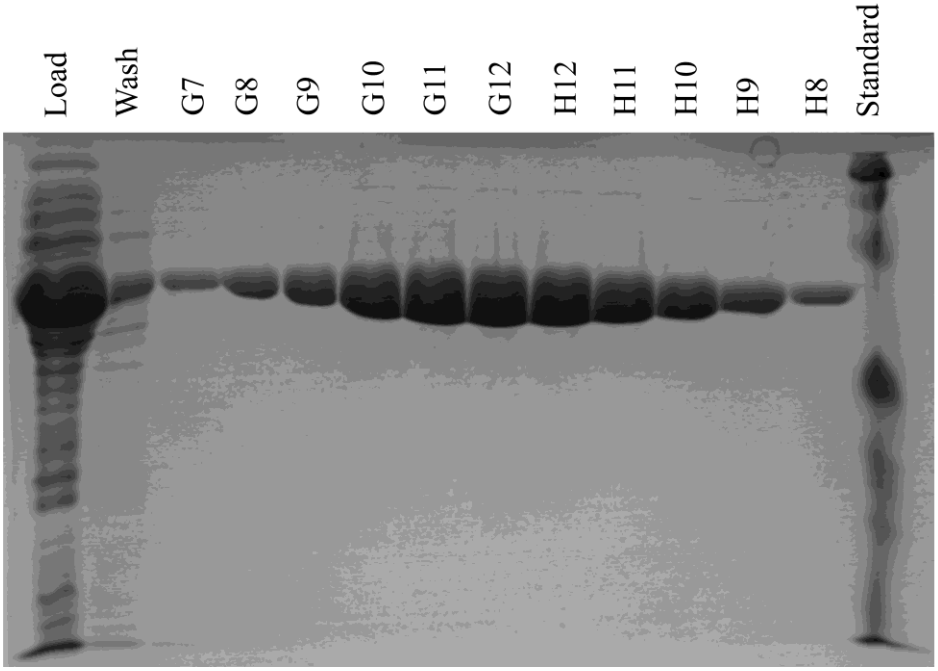

Figure 2.3 15% SDS-PAGE gel of *En. casseliflavus* GK fractions after the MonoQ column show highly purified protein.

2.2.6 GK-GlpO pull-down assay

Experience with attempting GK-HPr cocrystallizations, as well as initial pull-down and SPR work by Dr. Shoucheng Du and Dr. Haibin Shi suggested that association of GK and other members of the glycerol metabolism pathway was buffer-dependent. His-tagged *P. aeruginosa* GK was incubated overnight on cobalt resin beads so that the protein could bind to the cobalt beads. 1 mL syringes were prepared as mini columns by removing the plunger and putting a porous plastic plug into the bottom of the syringe. The GK-bound beads were aliquoted into 14 different 1 mL syringes that contained a plastic plug on the opening to prevent the beads from

leaking out of the syringe. The beads were washed with a buffer, each buffer containing slightly different conditions. The *Streptococcus pyogenes* GlpO was put over the beads, washed and eluted as in the purifications. This procedure allowed a much higher throughput than previous procedures, and did not require that the native gels be tailored specifically to the conditions tested. No buffer contained a pH lower than 6.0, as that would prevent the GK from interacting with the column. None of the buffers used had structures close to Tris (i.e. bicine, tricine, BTP, etc.).

2.2.7 *En casseliflavus* GK-*En. faecalis* HPr crystallization

Initial cocrystsallization attempts used *En. casseliflavus* GK and *En. faecalis* HPr incubated together in molar ratios of 1:1 to 1:5. (For further reference, this combination is listed as GK:HPr.) After incubation of 5-30 min, 96-well crystal screens were set up using 80 µL of mother liquor, 0.4-0.5 µL of protein between 8-15 mg/mL and 0.4-0.5 µL of crystallant.

Later attempts at cocrystallization used *En. casseliflavus* GK and *En. faecalis* HPr that were copurified together. 50 µL of HPr were neutated overnight at 4°C with a 50% slurry of 5 mL cobalt beads and 50 mM Tris-HCl pH 7.4, 30 mM sodium chloride, 10 mM imizadole. A 1.5 molar excess (1.25 mL of GK) was incubated overnight with cobalt in a 1:5 molar ratio of GK:cobalt, intended to bind to the His tags to prevent the GK from binding directly to the column.

Cobalt beads with *En. faecalis* HPr were put into a column and washed with 50 mL of 50 mM Tris-HCl pH 7.4, 10 mM imidazole, 300 mM sodium chloride, 10 mM imidazole. The pre-incubated GK was flowed over the cobalt beads, then washed with another 50 mL of the previous buffer. The beads were eluted with two sets of 10 mL (20 mL total) of 50 mM Tris-

HCl pH 7.4, 300 mM sodium chloride, 150 mM imidazole. Fractions were run on a gel to confirm copurification. The proteins were concentrated to 10 mg/mL and used for 96-well trays. Conditions which showed crystals or crystalline material were used to create larger drops in 24-well trays, with 500-800 µL mother liquor in the well and drops containing ratios of protein:crystallant between 1:4 and 4:1.

2.2.8 GK Crystallization

Initial *En. casseliflavus* GK was obtained from Dr. Ruth L. Saxl at 10-15 mg/mL. Protein was concentrated in an Amicon filter to 10-15 mg/mL prior to crystallization setup or diluted using buffer. The buffers were exchanged into 50 mM Tris pH 7.5 with or without 10% glycerol. I assisted Dr. Ruth L. Saxl with crystallization setups and crystal optimization.

Crystallization tests used protein concentrations between 8-20 mg/mL in 50 mM Tris pH 7.5 with or without 10% glycerol. All crystals were obtained via vapor diffusion. Crystallization screens were set up in 96-well trays using 80 µL of mother liquor, with 0.5 µL protein and 0.5 µL of crystallant in the drop. Initial hits were repeated in 24-well trays with 500-800 µL of mother liquor and drops containing between 1:4 and 4:1 ratios of protein:crystallant. Initial hits were optimized by varying the protein:crystallant ratio, concentration of precipitating reagent and pH, with the goal of increasing the size of the crystals and the diffraction resolution. His232Arg crystals that produced the highest resolution data images were obtained from 0.05 M KH_2PO_4, 17 % w/v PEG 8000 with 10% glycerol and soaked with 23% ethylene glycol. His232Glu crystals that produced the highest resolution data images were obtained from 0.1 M HEPES pH 6.5, 10 mM $ZnSO_4$, 25 % w/v PEG 550 MME with 10% glycerol, and 0.05 M

KH_2PO_4, 20% w/v PEG 800 soaked with 25% ethylene glycol. Data collection and processing are described in 2.2.10.

P. aeruginosa GK was thawed on ice and buffer exchanged into 50 mM Tris pH 8.0, 10% glycerol. Protein was concentrated to 8-12 mg/mL and used in crystal screens in 96-well trays, using the Honeybee robot to set them up. Tray setups contained 0.5 µL of protein and 0.5 µL of precipitant in the drops, and 80 µL of precipitant in the well. Crystals were seen within 24 hr. For conditions that had initial hits, the conditions were replicated in a larger format containing 500-800 µL of mother liquor in the well, and drops containing 1:4 to 4:1 ratios of protein:precipitant. Vapor diffusion trays were set up using an electronic pipette with aliquots in 0.5 µL steps, which gave more consistent results than the manual pipette. Crystals were obtained from these drops and flash cooled in liquid nitrogen before mounting at 100K. See Figure 2.4 to Figure 2.9 for crystal images.

2.2.9 Cryocooling conditions

X-ray radiation of protein crystals creates free radicals and damages the biological material in the protein crystals. If unattenuated, radiation damage in protein crystals will significantly limit the resolution and the number of exposures a crystal will be able to support before the internal components are destroyed. In order to reduce the amount of radiation damage that is perpetuated during the x-ray exposure process, protein crystals are cooled to a temperature around 100K (-180°C) using liquid nitrogen. Using the mother liquor, and increasing the concentration of glycerol to a final concentration between 20-30%, protected crystals that were crystallized in glycerol-containing conditions. Crystals that did not contain glycerol were cryoprotected in 20-30% ethylene glycol with mother liquor. Cryoprotection allowed each structural model to be

represented by a single crystal, allowing further determination of which conformational changes were a product of the crystallization and which conformational changes were indicative of a functional change.

2.2.10 Data collection and Processing

En. casseliflavus His232Glu with glycerol was collected at SLS X06A, 1080 images, total of 270°. *En. casseliflavus* His232Arg with glycerol was collected at SLS with a total of 1883 images, total of 470.75°. *En. casseliflavus* His232Arg and His232Glu with ethylene glycol was collected using a Rigaku Superbright FRE generator and Saturn 944 CCD detector. Data was processed using d*trek [102] or HKL2000 [103]. Data for the His232Arg with glycerol went to 1.73Å. Data for the His232Glu with glycerol went to 1.75Å. Data for the His232Arg with ethylene glycol went to 2.30Å, and data for the His232Glu with ethylene glycol went to 2.65Å. Data statistics for the datasets are found in Table 2.1.

P. aeruginosa glycerol-bound GK data was collected on the HTC using an FRE Superbright generator for glycerol data in slices of 1°, 150 images (150° collected total). Data for the *P. aeruginosa* ethylene glycol-containing model was collected at the GM/CA beamline at the Advanced Photon Source at Argonne National Laboratory in Chicago in slices of 1° rotation, for a total of 299 images (299° collected total). Ethylene glycol-bound *P. aeruginosa* GK data was processed in d*trek [102] to 1.75Å. Glycerol-bound *P. aeruginosa* GK data was processed by Dr. Ruth L Saxl to 1.73Å resolution in d*trek [102]. See Table 2.2 for data processing statistics.

GK:HPr crystals were collected on a Saturn 944 CCD using an FRE Superbright generator. Data was processed using d*trek [102]. Two forms of crystals were examined: one

form contains a $P2_12_12$ spacegroup similar to *En. casseliflavus* GK, and the second contains a P3 spacegroup. See Table 2.3 for statistics.

Table 2.1 Data processing statistics for *En. casseliflavus* His232Arg and His232Glu models

	His232Arg with glycerol	His232Arg with ethylene glycol	His232Glu with glycerol	His232Glu with ethylene glycol
Total reflections	930,616	335,073	891,523	181,650
Unique reflections	115,245	92,379	100,197	61,660
Space group	$P2_12_12$	$P2_1$	$P2_12_12$	$P2_1$
Unit cell parameters				
a (Å)	96.81	98.89	96.68	98.61
b (Å)	200.46	103.89	199.55	105.19
c (Å)	56.51	114.17	56.81	114.30
α (°)	90	90	90	90
β (°)	90	114.3	90	114.55
γ (°)	90	90	90	90
Resolution (Å)	50-1.73	50-2.3	50-1.75	50-2.65
R_{sym} (%)[ab]	8.7 (49.8)	8.2 (52.8)	9.5 (47.4)	8.3 (30.4)
I/(σ)I	9.3 (2.8)	11.1 (2.1)	8.4 (2.5)	7.5 (2.1)
Redundancy	8.08 (8.06)	3.8 (3.6)	8.9 (2.6)	2.95 (2.90)
Completeness (%)	99.7 (100)	98.5 (97.0)	97.6 (81.4)	99.7 (99.9)
Mosaicity	0.695	0.485-0.686	0.701	0.492
X^2	0.97 (1.02)	1.01 (1.06)	1.02	0.99 (1.31)
Number of images	1,080	180	1,080	322
Number of images used	802	180	1,068	141
Degrees per image (°)	0.25	1	0.25	1
Total oscillation (°)	270	180	270	322
Oscillation used (°)	200.5	180	267	141

[a] Numbers in parentheses are values for the highest resolution shell (out of ten shells).

[b] $R_{sym}=\Sigma_h\Sigma_i|I(h,i)-<I(h)>|/\Sigma_h\Sigma_i I(h,i)$ where $I(h,i)$ is the intensity values of the ith measurement of h and $<I(h)>$ is the corresponding mean values of $I(h)$ for all i measurements.

Table 2.2 Data processing statistics for the *P. aeruginosa* GK with either glycerol or ethylene glycol

	GK with glycerol	GK with ethylene glycol
Total Reflections	340,717	420,822
Unique Reflections	66,853	69,431
Space Group	I222	I222
a (Å)	62.31	62.55
b (Å)	115.15	114.81
c (Å)	190.03	190.58
α (°)	90	90
β (°)	90	90
γ (°)	90	90
Resolution (Å)	50-1.73	50-1.75
R_{sym} (%)[ab]	4.5 (48.0)	5.8 (53.5)
Completeness (%)	91.7 (54.9)	99.9 (99.9)
I/σ(I)	14.0 (2.1)	11.0 (2.4)
Redundancy	5.10 (2.95)	6.06 (5.96)
Mosaicity	1.07	0.75
X^2	0.95 (1.25)	0.95 (1.26)
Total Number of Images	299	150
Rotation per Image	0.5	1
Total Rotation	149.5	150
Number of Images Used	288	150
Rotation Used	144	150

[a] Numbers in parentheses are values for the highest resolution shell (out of ten shells).

[b] $R_{sym}=\Sigma_h\Sigma_i|I(h,i)-<I(h)>|/\Sigma_h\Sigma_i I(h,i)$ where $I(h,i)$ is the intensity values of the ith measurement of h and $<I(h)>$ is the corresponding mean values of $I(h)$ for all i measurements.

Table 2.3 Data processing statistics for representative *En. casseliflavus* GK-*En. faecalis* HPr cocrystallization trials

	GK-HPr 1	GK-HPr 2
Total Reflections	203,740	335,362
Unique Reflections	102,995	68,241
Resolution (Å)	50-3.10	50-2.05Å
Space group	P3	$P2_12_12$
a (Å)	161.3	95.893
b (Å)	161.3	200.851
c (Å)	194.5	56.281
α (°)	90	90
β (°)	90	90
γ (°)	120	90
Completeness (%) [a]	100 (100)	15.2 (44.1)
I/σ(I)	8.09 (1.74)	8.7 (4.1)
R_{sym} (%) [b]	18.1 (47.2)	4.9 (4.9)
Redundancy	7.9 (7.2)	98.0 (98.6)

[a] Numbers in parentheses are values for the highest resolution shell (out of ten shells).
[b] $R_{sym}=\Sigma_h\Sigma_i|I(h,i)-<I(h)>|/\Sigma_h\Sigma_i I(h,i)$ where $I(h,i)$ is the intensity values of the ith measurement of h and $<I(h)>$ is the corresponding mean values of $I(h)$ for all i measurements.

2.2.11 Molecular replacement

The *En. casseliflavus* His232Ala model determined by Dr. Joanne Yeh, et al (PDB ID 3D7E) was modified to contain only protein from residue 5 to residue 499. Molecular replacement was done in PHASER [104] using a script. The script allowed PHASER to search in multiple point groups (command SGALTERNATIVE all) to determine the screw axes. If a screw axis is present, certain periodic reflections will be missing, caused by the asymmetric unit being shifted by a fraction of the unit cell (usually shifted by half). Combining this information with the missing reflections allowed for confirmation of the space group. The rotational and translational Z values for each monomer are in Table 2.4.

Molecular replacement in PHASER (as a subprogram of PHENIX) for the *P. aeruginosa* GK with glycerol utilized the protein content of *E. coli* GK model PDB ID 1GLA. Refinement of the glycerol-bound protein was partially completed before using the *P. aeruginosa* GK as a model for molecular replacement in the ethylene glycol-bound structure. Molecular replacement statistics are found in Table 2.5.

2.2.12 Model refinement

All structures were refined in CNS [105], CCP4 [106] and PHENIX [107] using alternating refinement and manual model fitting in Coot [108]. The R_{work} and R_{free} of each structure and data refinement statistics are listed in Table 2.4 and Table 2.5. Structures from GK:HPr crystals do not contain waters or ligands.

Table 2.4 Molecular replacement and refinement statistics for the *En. casseliflavus* His232ArgR and His232Glu models with glycerol or ethylene glycol.

	His232Arg with glycerol	His232Arg with ethylene glycol	His232Glu with glycerol	His232Glu with ethylene glycol
Molecular replacement				
Rotation Z	24.0	21.1	30.6	23.9
Translation Z	33.4	20.4	41.0	18.6
Rotation Z	29.3	21.3	24.0	27.3
Translation Z	89.5	59.2	27.6	59.3
Rotation Z		20.4		23.4
Translation Z		67.6		69.1
Rotation Z		22.9		27.0
Translation Z		82.3		87.5
Clashes	0	0	0	0
R_{work} (%)		30.95		23.94
R_{free} (%)		31.51		23.62
Refinement statistics				
R_{work} (%)	20.74	18.14	19.38	20.80
R_{free} (%)	21.61	20.56	22.79	25.13
Bond angle RMSD (°)	1.27	1.33	1.28	1.36
Bond length RMSD (Å)	0.005	0.006	0.006	0.007
Ramachandran statistics				
Most favored (%)	91.1	82.0	90.0	86.0
Allowed (%)	8.1	16.2	9.7	12.5
Generously allowed (%)	0.3	1.4	0.1	0.8
Disallowed (%)	0.5	0.3	0.2	0.7

Table 2.5 Molecular replacement and refinement statistics for *P. aeruginosa* GK with glycerol or ethylene glycol

	PAGK with glycerol	PAGK with ethylene glycol
Molecular replacement		
Rotation Z	32.3	44.0
Translation Z	61.7	54.0
Clashes	0	0
R_{work} (%)	--	30.59
R_{free} (%)	--	30.44
Refinement statistics		
R_{work} (%)	20.02	17.56
R_{free} (%)	22.93	20.04
Bond angle RMSD (°)	0.99	1.07
Bond length RMSD (Å)	0.007	0.007
Ramachandran statistics		
Most favored (%)	95.7	93.5
Allowed (%)	3.3	5.2
Generously allowed (%)		
Disallowed (%)	1.0	1.3
Rotamer outliers (%)	3.6	2.5
C_β outliers	0	0
Clashscore	14.81	20.50

2.2.13 Model analysis

Coot was used to manually examine structures and calculate distances between specific atoms and areas of the model [109]. DDMP was used to determine the relationships between residues

of the different models. It calculates the distances between all of the Cα atoms in a model, and compares those distances to the distances of a second model [110]. This comparison allows detection of areas that are shifted relative to other portions of the model and displays results in a format that is easily readable. Data was examined using Coot [109], CNS [105] and DDMP [110]. Superpositions for comparisons were done in either Coot (for the entire molecule or monomer) [109] or CCP4 [106] for superimposing only a portion of the model. In cases where only a portion of the model was superimposed, portions of the model that had a high amount of flexibility or variability between the models were omitted from the superposition (the activation loop and the termini). Coot or CCP4 was used to determine distances between atoms. DDMP was used to determine which portions of the models had significant movement relative to other portions of the model. PDB files were modified to omit any secondary conformations or ligands and have a consistant number of residues in each model.

LSQMAN from the DEJAVU package from Uppsala [111, 112] was used to determine the RMSD and rotation necessary to superimpose the *P. aeruginosa* tetramer on the *E. coli* models. Residues 5 to 490 of the *E. coli* structures were superimposed onto residues 8 to 498 of the *P. aeruginosa* structures to obtain the RMSD of the O monomer, then comparable residues were superimposed for the X, Y and Z monomers to obtain the rotations and RMSD of the other monomers in the tetramer.

2.3 RESULTS

2.3.1 Crystal description

Crystals of the glycerol-bound *En. casseliflavus* His232Arg were obtained from modified Hampton crystal screen I 42 with 0.05 M KH_2PO_4, 17% PEG 8000 and 10% glycerol (see Figure 2.4). Crystals of the ethylene glycol-bound *En. casseliflavus* His232Arg and His232Glu were obtained from modified Hampton crystal screen I 42 with 0.05 M KH_2PO_4, 17-20% PEG 8000 and cryoprotected in ethylene glycol (see Figure 2.5 and Figure 2.6). Crystals of the glycerol-bound *En. casseliflavus* His232Glu were obtained from in Hampton crystal screen II 27, 0.01 M $ZnSO_4$, 0.2 M MES pH 6.5, 25% v/v PEG 550 MME and 10% glycerol (see Figure 2.7). Crystals were mostly tetrahedral in shape, and approximately 0.05 to 0.2 mm in size along the longest axis. Crystals grown in glycerol were sturdier and easier to loop and cool, whereas ethylene glycol-bound crystals had a tendency to crack during cryoprotection.

Crystals of *P. aeruginosa* were grown in a variety of conditions with the same I222 space group, suggesting that the crystal packing is very stable. Crystals were clear, but strongly dichroic. Crystals used to determine the structure of *P. aeruginosa* with glycerol came from Hampton cryo screen II 15, consisting of 0.375 M ammonium sulfate, 0.075 M sodium citrate tribasic dehydrate pH 5.6, 0.75 M lithium sulfate and 25% v/v glycerol (see Figure 2.8). The glycerol-bound crystal was further cryoprotected in mother liquor with ethylene glycol. *P.aeruginosa* crystals with ethylene glycol were crystallized in Emerald Wizard I 8, consisting of 2.0 M ammonium sulfate, 0.1 M sodium citrate/citric acid and cryoprotected with 25% ethylene glycol (see Figure 2.9). Crystals were large, between 0.1 to 0.4 mm, and easily looped and flash cooled.

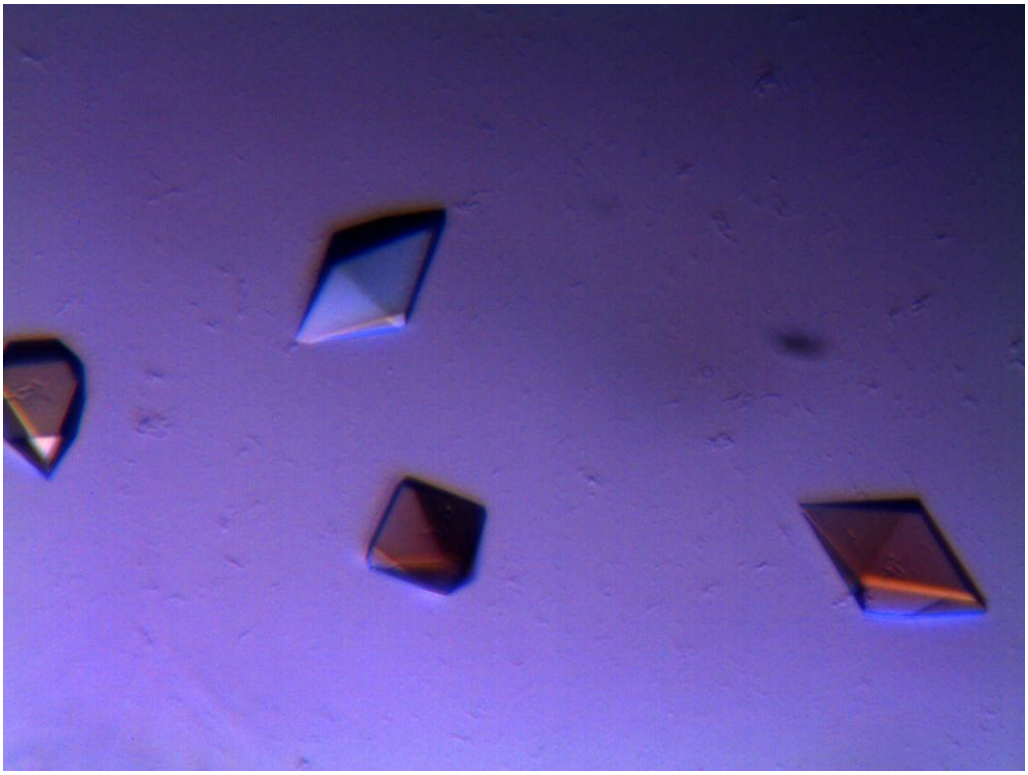

Figure 2.4 *En. casseliflavus* His232Arg mutant with glycerol from a modified Hampton crystal screen I 42 with 0.05 M KH_2PO_4, 17% PEG 8000 and 10% glycerol. Image taken by Dr. Ruth L. Saxl.

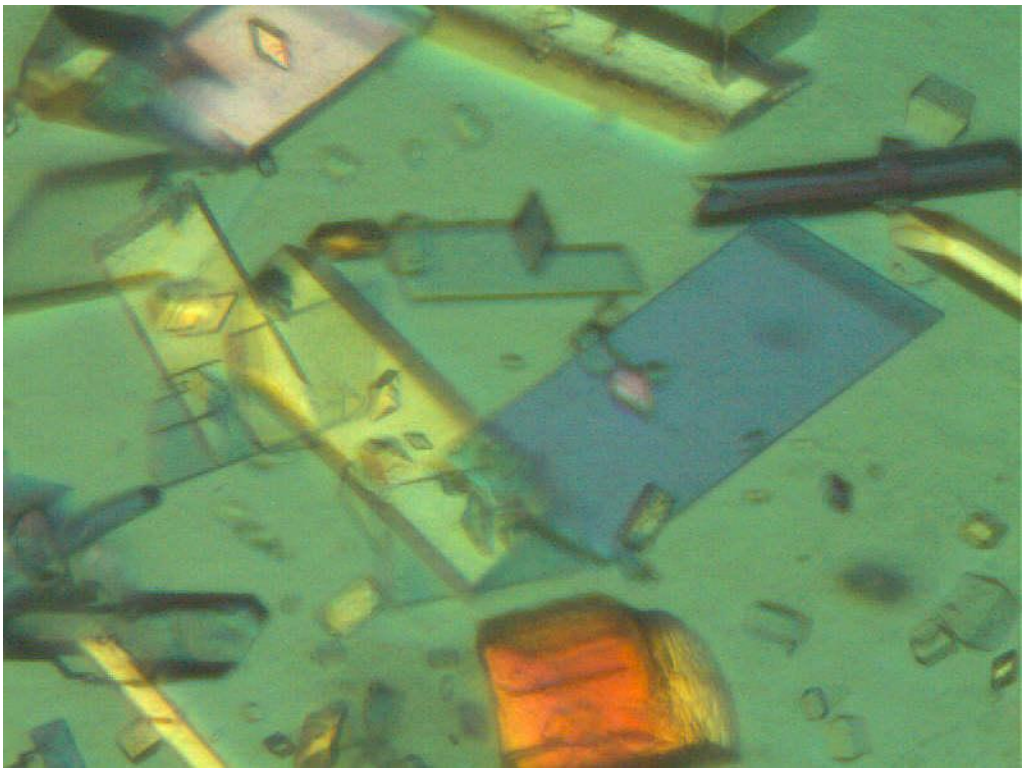

Figure 2.5 *En. casseliflavus* His232Arg mutant with ethylene glycol in Hampton crystal screen I 42, 0.05 M KH_2PO_4 and 17% PEG 8000. Image taken by Dr. Ruth L. Saxl.

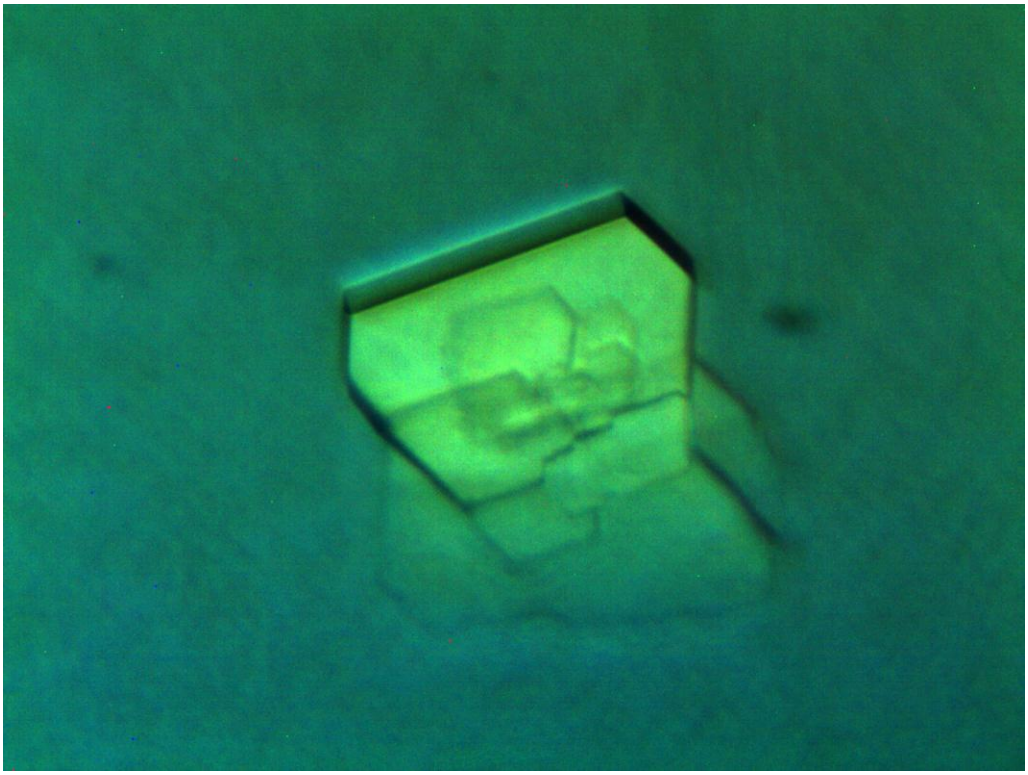

Figure 2.6 ***En. casseliflavus*** **His232Glu mutant with ethylene glycol in Hampton crystal screen I 42, 0.05 M KH_2PO_4 and 20% PEG 8000. Image taken by Dr. Ruth L. Saxl.**

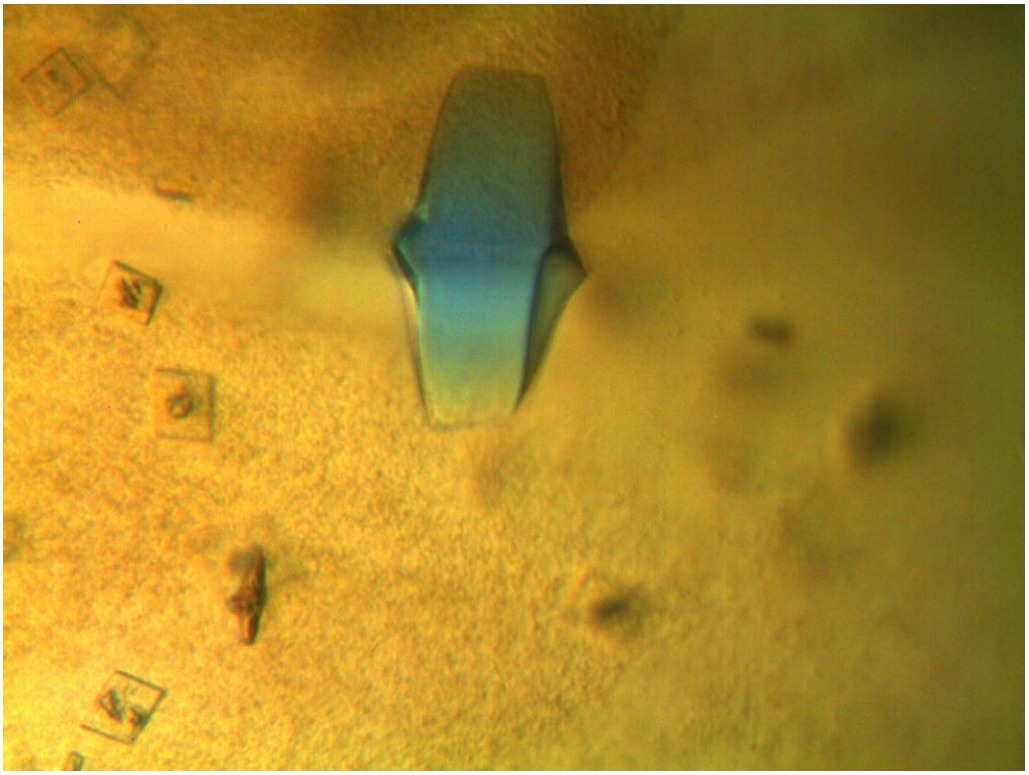

Figure 2.7 *En. casseliflavus* His232Glu with glycerol in Hampton crystal screen II 27, 0.01 M $ZnSO_4$, 0.2 M MES pH 6.5 and 25% v/v PEG 550 MME. Image taken by Dr. Ruth L. Saxl.

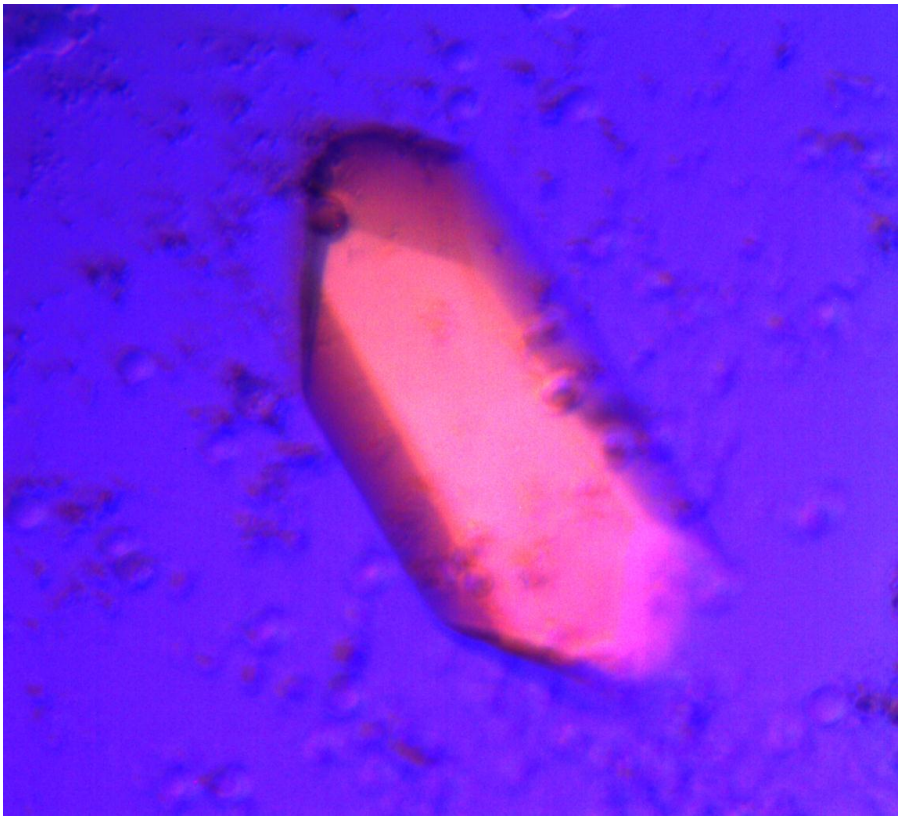

Figure 2.8 *P. aeruginosa* GK with glycerol in Hampton cryo II 15, consisting of 0.375 M ammonium sulfate, 0.075 M sodium citrate pH 5.6, 0.75 M lithium sulfate and 25% v/v glycerol. Image taken by Dr. Ruth L. Saxl.

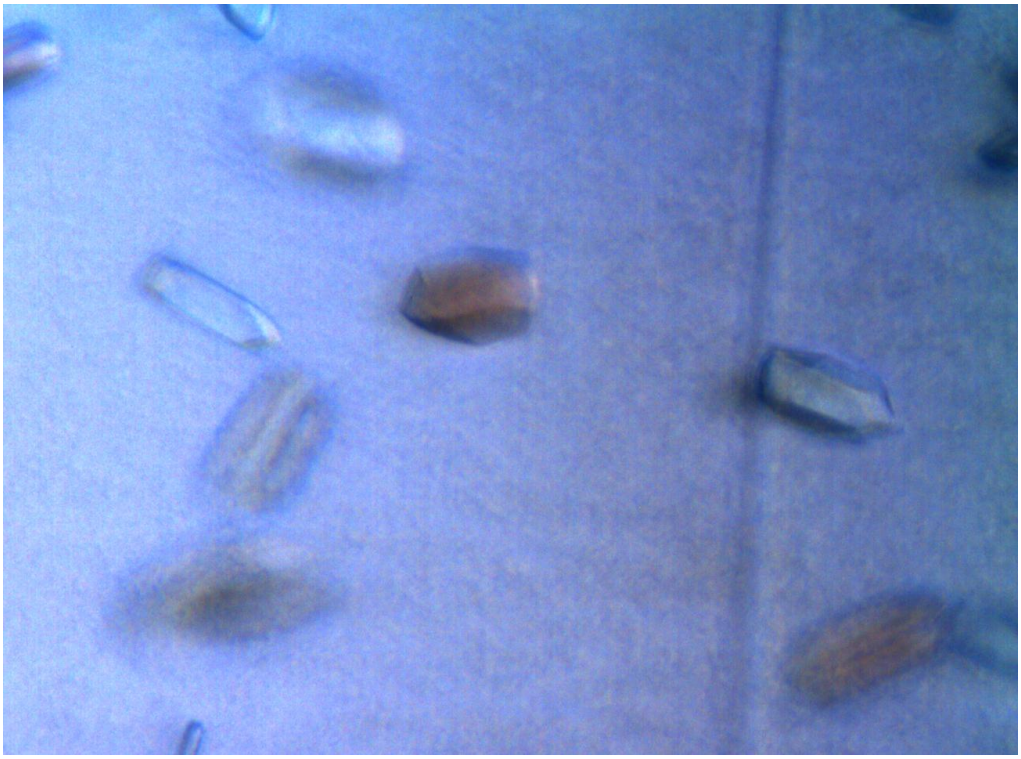

Figure 2.9 *P. aeruginosa* GK with ethylene glycol crystallized in Wizard I 8, consisting of 2.0 M ammonium sulfate, 0.1 M sodium citrate pH 5.5. Image taken by Dr. Ruth L. Saxl.

2.3.2 Crystallization of GK-HPr

Crystallization trials of GK:HPr gave crystals that had two distinct morphologies: they either looked very similar to the previous *En. casseliflavus* GK crystals, or had a hexagonal shape to them (see Figure 2.4 and Figure 2.10). The crystals that had the same morphology as previous crystals also had unit cells and space groups that were very similar to the previous crystals of pure GK monomers (compare the data statistics of Table 2.3 with the data statistics of Table 2.1). Hexagonal crystals contained a different orientation of GK in the asymmetric unit. The asymmetric unit contained a full tetramer of GK oriented in a triangular orientation, rather than the P2-type orientation found in previous WT and mutant GK crystals.

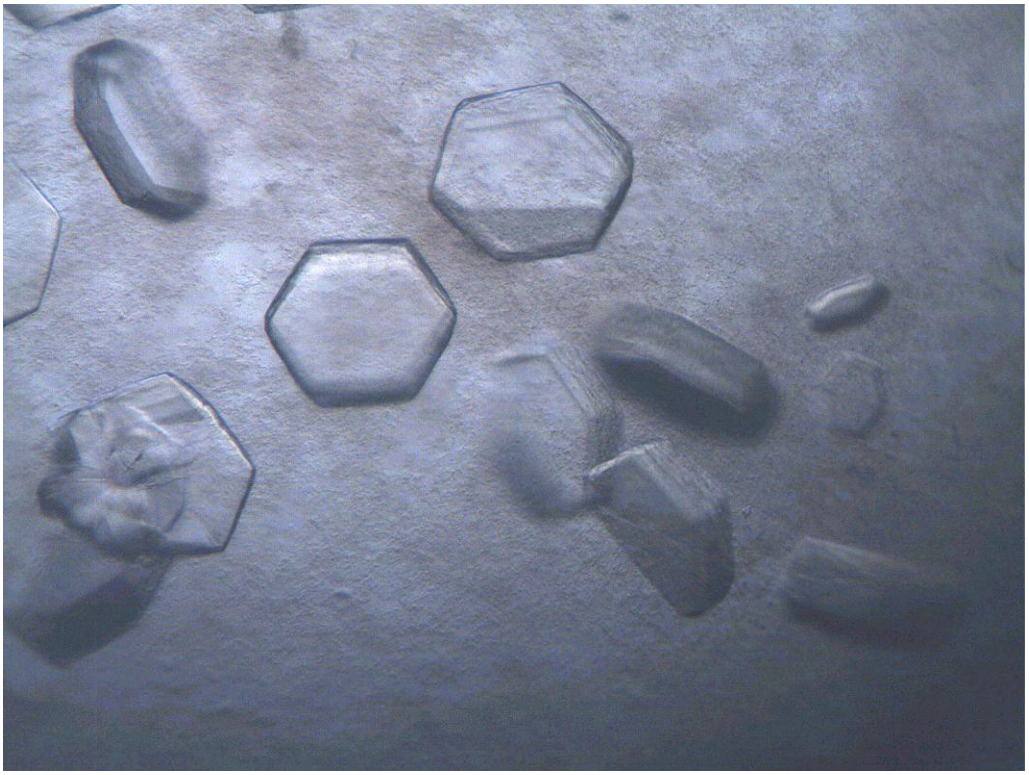

Figure 2.10 Crystals of GK:HPr in Emerald cryo I 41, consisting of 25% v/v 1,2-propanediol, 0.1 M sodium phosphate/citrate pH 4.2, 5% w/v PEG 3000, 10% v/v glycerol.

2.3.3 Overall tetrameric structure

For all of the *En. casseliflavus* and *P. aeruginosa* GK models, an OXYZ tetramer was formed in the crystal (see Figure 1.2, nomenclature as in [54]). The *En. casseliflavus* glycerol-bound models contained the OX dimer in the asymmetric unit; the tetramer could be created by using the correct symmetry operations. The *En. casseliflavus* models that contain ethylene glycol present in the active site contain the OXYZ tetramer in the asymmetric unit. In the glycerol-bound tetramer, the OX dimer is identical to the YZ dimer, but when ethylene glycol is present there are slight differences between the OX and YZ dimers, namely the ethylene glycol

orientation. The *P. aeruginosa* GK models contained one monomer in the asymmetric unit, denoted as O. The tetramer could be obtained through using the appropriate symmetry-related monomers.

The *P. aeruginosa* GK tetramer was compared to previous *E. coli* models to determine whether the model showed a physiologically active or inactive model. Table 2.6 contains a list of models used for comparison, along with bound ligands and select structure characteristics. Table 2.7, Table 2.8 and Table 2.9 contain the angles necessary to superimpose the O, Y, and Z monomers of each of the tetramers once the X monomer has been superimposed. This shows that the models are closest to 1BU6 [113]. Although the rotation necessary to superimpose the O is smaller for 1BO5, the larger rotation necessary to superimpose the Z and Y monomers suggests a greater overall difference between the *P. aeruginosa* GK models and the structure of PDB ID 1BO5 [53].

Table 2.6 *Escherichia coli* models currently in the PDB. The PDB ID, select molecular statistics, mutations, and other ligands are listed for reference.

PDB ID	Resolution (Å)	R values (%)	Ligands	Mutation
1BWF [57]	3.00	16.8	GOL, ATF, Mg	S58W
3EZW [114]	2.00	16.8 (22.5)	GOL, EDO, Cl, Mg	G230D
1GLJ [57]	3.00	16.6	ATS, GOL, Mg	S58W
1GLL [57]	3.00	17.6	ACP, GOL, Mg	S58W
1BOT [53]	3.05	21.9	EPE, GOL	--
1BO5 [53]	3.20	21.1	FBP, GOL	--
1GLF [113]	2.62	14.6	ADP, GOL, PO4	--
1BU6 [113]	2.37	16.7	GOL, SO4	none listed, A65T in print
1GLA [54]	2.60	19.1	GOL	$EIIA^{Glc}$
1GLB [54]	2.60	20.5	GOL, ADP	$EIIA^{Glc}$

Table 2.7 Comparison of *P. aeruginosa* GK tetramer with active and inactive *E. coli* GK tetramers

	PA GK with glycerol				PA GK with ethylene glycol			
	D I RMSD[a]	D II RMSD[a]	angle[b]	DII RMSD[c]	D I RMSD[a]	D II RMSD[a]	angle[b]	DII RMSD[c]
3EZW [114]	1.1	0.6	6.2	2.5	1.1	0.7	6.8	2.7
1BWF [57]	1.1	0.6	4.9	2.3	1.1	0.6	5.4	2.6
1BO5 [53]	1.1	0.7	8.7	2.3	1.1	0.8	8.9	2.7
1GLF [113]	1.1	0.7	6.0	2.5	1.1	0.8	6.6	2.8
1BU6 [113]	1.1	0.7	4.0	2.3	1.1	0.7	4.3	2.6
1GLA [54]	1.2	0.7	6.8	2.4	1.2	0.7	7.1	2.7

[a] RMSD was calculated for Cα atoms of Domain I residues 8-227, 242-260 and 442-461 or Cα atoms of Domain II residues 265-277, 306-310 and 341-435 after superposition.
[b] The angle required to superimpose Domain II after superposition of Domain I.
[c] RMSD was calculated from the PA GK Domain II Cα of residues 265-277, 306-310 and 341-435 after superposition of the Domain I Cα of residues 8-227, 242-260 and 442-461.

Table 2.8.Rotation necessary to superimpose other monomers onto *P. aeruginosa* GK with glycerol

	X RMSD (Å)	O Angle (°)	Z Angle (°)	Y Angle (°)
1BO5 [53]	1.7	3.2	8.2	8.2
1BU6 [113]	1.7	5.0	5.9	5.8
1BWF [57]	1.3	12.1	5.0	11.4
3EZW [114]	1.6	9.8	12.8	4.8
1GLF [113]	1.7	4.7	4.7	4.7
1GLA [54]	1.3	3.5	3.7	4.0

Table 2.9. Rotation necessary to superimpose other monomers onto *P. aeruginosa* GK with ethylene glycol

	X RMSD (Å)	O Angle (°)	Z Angle (°)	Y Angle (°)
1BO5 [53]	1.7	2.6	11.1	10.9
1BU6 [113]	1.8	2.7	8.4	8.8
1BWF [57]	1.5	11.2	6.8	10.5
3EZW [114]	1.7	10.7	13.7	7.3
1GLF [113]	1.8	2.6	7.0	7.6
1GLA [54]	1.4	2.4	2.5	2.8

2.3.4 Glycerol and ethylene glycol interactions

In all glycerol-bound models, each monomer contains one glycerol in the active site. In the *En. casseliflavus* structures the glycerol is aligned such that the hydroxyl groups hydrogen bound with Arg84, Tyr136, Asp246, Gln247, Glu85, and the carbons form a hydrophobic surface with the Trp104 and Phe271 (see Figure 2.11). The image shows the residues that contribute to the hydrogen bonds or charged interactions in blue, while hydrophobic residues are shown in yellow. The presence of a hydrophobic and hydrogen bonding region stabilizes and orients the glycerol in the active site. The interactions between the glycerol and active site residues are similar regardless of mutant (see Table 2.10 to Table 2.12). All *En. casseliflavus* GK structures contain a phosphate or sulfate ion present near the active site. Comparison of *E. coli* structures shows the ion to be positioned where the gamma phosphate of ATP would be present.

The *P. aeruginosa* glycerol-bound model shows glycerol hydrogen bonding with Arg87, Tyr139, Asp249, Gln250 and Glu88. The glycerol also interacts with the hydrophobic surface created by Trp107 and Phe249.

The models that contain ethylene glycol show one molecule of ethylene glycol in the active site in a conformation very similar to the glycerol in the glycerol-bound model. In the *En. casseliflavus* models the ethylene glycol hydrogen bonds with Arg84, Tyr136, Glu85, and Asp 246. The hydrophobic pocket created by Phe271 and Trp104 holds the two carbons (see Figure 2.12). Because the third carbon and hydroxyl is not present in ethylene glycol as it is in glycerol, the hydrogen bonds created by those atoms are not present in the ethylene glycol models. The *P. aeruginosa* model only contains one monomer per asymmetric unit, so only one set of interactions is observed, rather than the variants that are observed in the *En. casseliflavus* model. The ethylene glycol hydrogen bonds with residues Arg87, Tyr139, Asp249 and Glu88, with a

hydrophobic interaction with Trp107 and Phe274. The interactions between the glycerol and glycerol kinase closest to the entrance of the catalytic cleft help to close the glycerol kinase active site. The ethylene glycol mutants show variability in the contacts created between the ethylene glycol and the GK mutants. Table 2.13 to Table 2.15 list the hydrogen bonding distances between the *En. casseliflavus* GK active site residues and the ethylene glycol. Each monomer in the asymmetric unit uses a slightly different hydrogen bonding distance between the active site residues and the O2. The hydroxyl group furthest from the active site opening appears to be held by the same interactions regardless in the glycerol-bound and ethylene glycol-bound structures. The conformation of the ethylene glycol makes the interactions between the O2 and the protein to be more inconsistent. This affects the closure of the active site (see 2.3.6 for more details).

The glycerol molecule is oriented so that C1 and O1 are furthest back in the active site, away from the active cleft opening, and C3 and O3 are closest to the active cleft opening where the γ phosphate of ATP would be placed. The residues that interact with glycerol in the mutants are the same residues found to interact with glycerol in the WT structure. As is noted in Table 2.12, the hydrogen bonding distances between glycerol and the active site residues are not significantly affected by the mutations, producing hydrogen bonding distances that are consistent between the models.

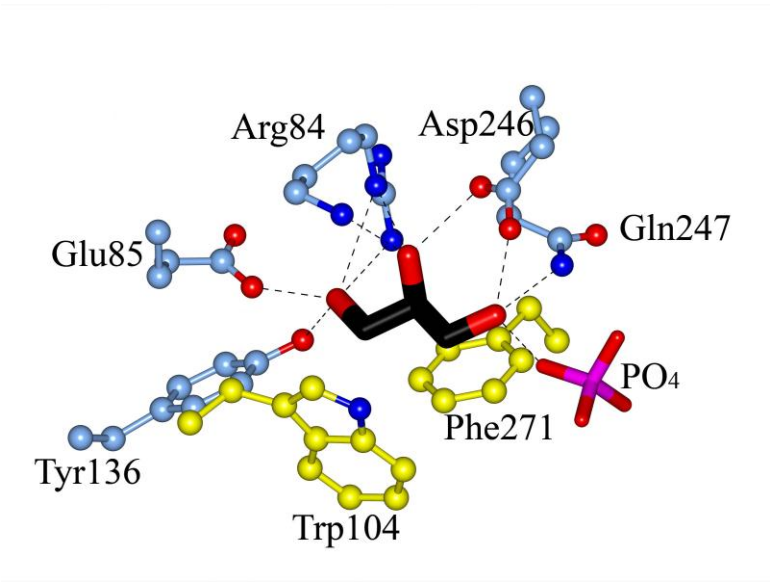

Figure 2.11 *En. casseliflavus* glycerol-bound active site Hydrophobic residues are shown in yellow, while hydrophilic residues are shown in blue. The phosphate in magenta is at the entrance to the active cleft.

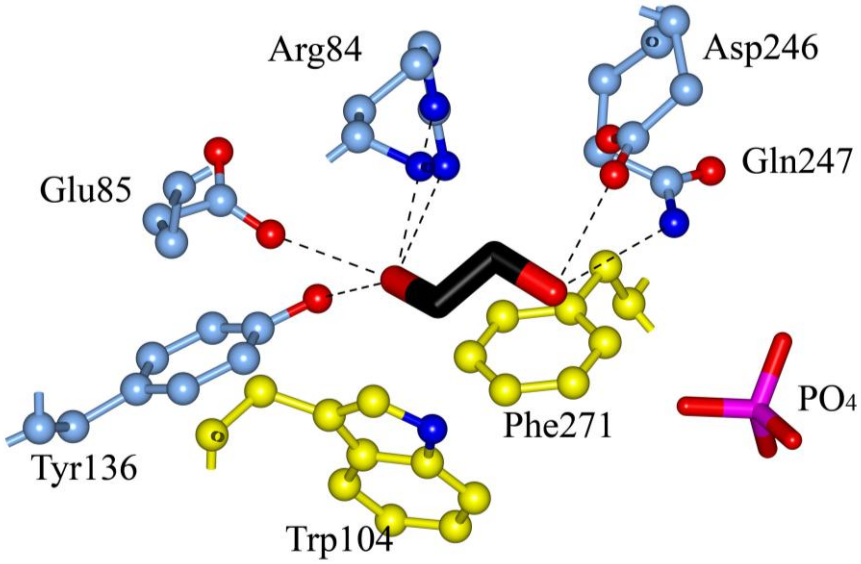

Figure 2.12 *En. casseliflavus* ethylene glycol active site. Hydrophobic residues are shown in yellow, while hydrophilic residues are shown in blue. The phosphate in magenta is at the entrance to the active cleft.

Table 2.16 lists the hydrogen bonding distances between the *P. aeruginosa* active site residues and the bound glycerol or ethylene glycol. Although the ethylene glycol-bound structure does not show any variation in the interactions between the ethylene glycol and the active site residues, the O1 atom is anchored as in the *En. casseliflavus* GK and the interactions with residues closest to the active cleft opening are absent.

Table 2.10 *En. casseliflavus* His232Arg hydrogen bonding distances between glycerol and active site residues

Glycerol Kinase Atom (Å)	Glycerol Atom	Distance in Monomer O (Å)	Distance in Monomer X (Å)
Glu85 OE1	O1	2.5	2.5
Glu85 OE2	O1	3.3	3.3
Tyr136 OH	O1	2.7	2.8
Arg84 NE	O1	3.2	3.2
Arg84 NH	O1	3.0	3.0
Arg84 NE	O2	3.0	3.0
Arg84 N	O2	3.0	2.9
Asp246 OD1	O2	2.6	2.7
Asp246 OD2	O2	3.4	3.4
Asp246 OD1	O3	3.4	3.5
Asp246 OD2	O3	2.5	2.7
Gln247 NE2	O3	3.1	3.1

Table 2.11 *En. casseliflavus* His232Glu hydrogen bonding distances between the glycerol and glycerol kinase active site residues

Glycerol Kinase Atom (Å)	Glycerol Atom	Distance in Monomer O (Å)	Distance in Monomer X (Å)
Glu85 OE1	O1	2.6	2.7
Glu85 OE2	O1	3.3	3.5
Tyr136 OH	O1	2.6	2.6
Arg84 NE	O1	3.3	3.3
Arg84 NH	O1	2.8	2.8
Arg84 NE	O2	3.0	3.0
Arg84 N	O2	2.9	2.9
Asp246 OD1	O2	2.6	2.6
Asp246 OD2	O2	3.3	3.6
Asp246 OD1	O3	3.3	3.3
Asp246 OD2	O3	2.6	2.7
Gln247 NE2	O3	3.0	3.0

Table 2.12 Average hydrogen bonding distances (Å) between glycerol and glycerol kinase active site residues in *En. casseliflavus* mutants.

Glycerol Kinase Atom	Glycerol Atom	His232Arg Distance (Å)	His232Glu Distance (Å)
Glu85 OE1	O1	2.5	2.6
Glu85 OE2	O1	3.3	3.4
Tyr136 OH	O1	2.7	2.6
Arg84 NE	O1	3.2	3.3
Arg84 NH	O1	3.0	2.8
Arg84 NE	O2	3.0	3.0
Arg84 N	O2	3.0	2.9
Asp246 OD1	O2	2.7	2.6
Asp246 OD2	O2	3.4	3.5
Asp246 OD1	O3	3.5	3.3
Asp246 OD2	O3	2.6	2.7
Gln247 NE2	O3	3.1	3.0

Table 2.13 *En. casseliflavus* His232Arg hydrogen bonding distances between ethylene glycol and glycerol kinase active site residues.

Glycerol Kinase Residue and Atom	Ethylene Glycol Atom	Distance in Monomer O (Å)	Distance in Monomer X (Å)	Distance in Monomer Y (Å)	Distance in Monomer Z (Å)
Glu85 OE1	O1	2.7	3.0	2.7	3.0
Glu85 OE2	O1	3.6	---	3.9	---
Tyr136 OH	O1	2.7	2.6	2.8	2.8
Arg84 NE	O1	3.4	3.7	3.6	3.3
Arg84 NH	O1	3.2	3.3	3.5	3.1
Asp246 OE1	O2	3.0	3.2	2.9	2.8
Asp246 OE2	O2	---	3.4	3.6	3.4
Gln247 NE2	O2	3.9	3.7	3.8	3.8

Table 2.14 *En. casseliflavus* His232Glu hydrogen bonding distances between ethylene glycol and glycerol kinase active site residues

Glycerol Kinase Residue and Atom	Ethylene Glycol Atom	Distance in Monomer O (Å)	Distance in Monomer X (Å)	Distance in Monomer Y (Å)	Distance in Monomer Z (Å)
Glu85 OE1	O1	2.8	2.9	2.6	3.0
Glu85 OE2	O1	---	---	---	---
Tyr136 OH	O1	2.7	3.1	2.8	2.8
Arg84 NE	O1	3.3	3.6	3.3	3.8
Arg84 NH	O1	3.0	3.4	3.1	3.4
Asp246 OE1	O2	3.1	3.1	2.9	3.0
Asp246 OE2	O2	3.7	3.6	3.5	3.5
Gln247 NE2	O2	3.6	3.3	---	3.7

Table 2.15 Average hydrogen bonding distances between ethylene glycol and *En. casseliflavus* glycerol kinase active site residues

Glycerol Kinase Residue and Atom	Ethylene Glycol Atom	His232Arg Distance (Å)	His232Glu Distance (Å)
Glu85 OE1	O1	2.8 ±0.2	2.8±0.2
Glu85 OE2	O1	3.7±0.4	---
Tyr136 OH	O1	2.7±0.1	2.9±0.2
Arg84 NE	O1	3.5±0.2	3.5±0.3
Arg84 NH	O1	3.3±0.2	3.2±0.2
Asp246 OE1	O2	3.0±0.2	3.0±0.2
Asp246 OE2	O2	3.5±0.1	3.6±0.1
Gln247 NE2	O2	3.8±0.1	3.5±0.2

Table 2.16. Hydrogen bonding distances between *P. aeruginosa* active site residues and glycerol (GOL) or ethylene glycol (EDO).

GK Residue and atom	Atom (GOL[a])	Distance (Å)	Atom (EDO)	Distance (Å)
Tyr133 OH	O1	2.6	O1	2.6
Glu88 OE1	O1	2.5	O1	2.7
Glu88 OE1	O1	3.4	O1	3.5
Arg87 NH2	O1	3.1	O1	2.9
Arg87 NE	O1	3.4	O1	3.4
Arg87 NE	O2	3.0	O2	3.0
Arg87 O	O2	3.0	O2	2.7
Asp249 OD1	O2	2.6	O2	2.7
Asp249 OD2	O2	3.4	--	
Asp249 OD2	O3	2.7	--	
Asn250 NE2	O3	3.2	--	
H_2O	O3	2.9	--	

[a] The O1 atom is the oxygen atom furthest inside the active site cleft. Numbering proceeds incrementally to the entrance of the activation cleft.

2.3.5 *En. casseliflavus* activation loop

The activation loops of the His232Arg and His232Glu mutants show significant differences (see Figure 2.13). The His232Arg activation loop contains interactions between Arg232 and Asp176 that stabilizes the activation loop. The electron density for the His232Arg loop is much more apparent than for the His232Glu electron density, also reflected in the B factors. B factors give information about the relative movement of regions of the model. Higher B values demonstrate that the particular region of the structure samples more conformations than other regions of the model.

Arg232 has a charged interaction with Asp176. In the phosphorylated activated GK, the phosphorylated His232 would have the opposite charge, being negative instead of positive. This

would be repulsed by the Asp176 if the Asp is the only amino acid capable of interacting with the His232. However, there is a lysine within the same area that would be positively charged. This lysine would be close enough to the activation loop to interact with the negatively-charged phosphorylated His232, creating the same stabilization effects that would occur in the His232Arg mutant. Figure 2.8 shows the differences between the activation loop position caused by this stabilization. The activation loop of mutants with higher activity are positioned further into the space between the X and O dimer. Comparison between the His232Glu and His232Arg mutants show that the small changes in the activation loop result in a 2° closure of the binding site cleft. The closure also rotates the OX dimer so that the distance between Domain II in the OX dimer closes by 6 Å at the point furthest from the activation loop.

In the His232Glu mutant, the distance difference in amino acid length between the phosphorylated His232 and the mutated Glu232 is large enough that it is unable to interact with anything on Domain I near the activation loop. The distance from the Cα to the charged phosphate at the WT His232 is larger than the distance from the Cα to the charge in the mutated His232Glu. The distance between the Cα and the phosphate in the phosphorylated histidine is between 6-7Å, depending on the orientation of the histidine and whether you calculate from the phosphorus or the oxygen atom. However, the distance between the Cα and the oxygen atoms of glutamate are between 3.5-5Å, depending on the geometry, a much shorter distance. This distance is enough to prevent interaction with critical residues on Domain I, rendering the protein inactive. Without stabilization of the activation loop, the large aromatic residues in the activation loop become very disordered due to steric clashes. The clashes actually increase the distance between the O and X monomers, leading to a decrease in activity.

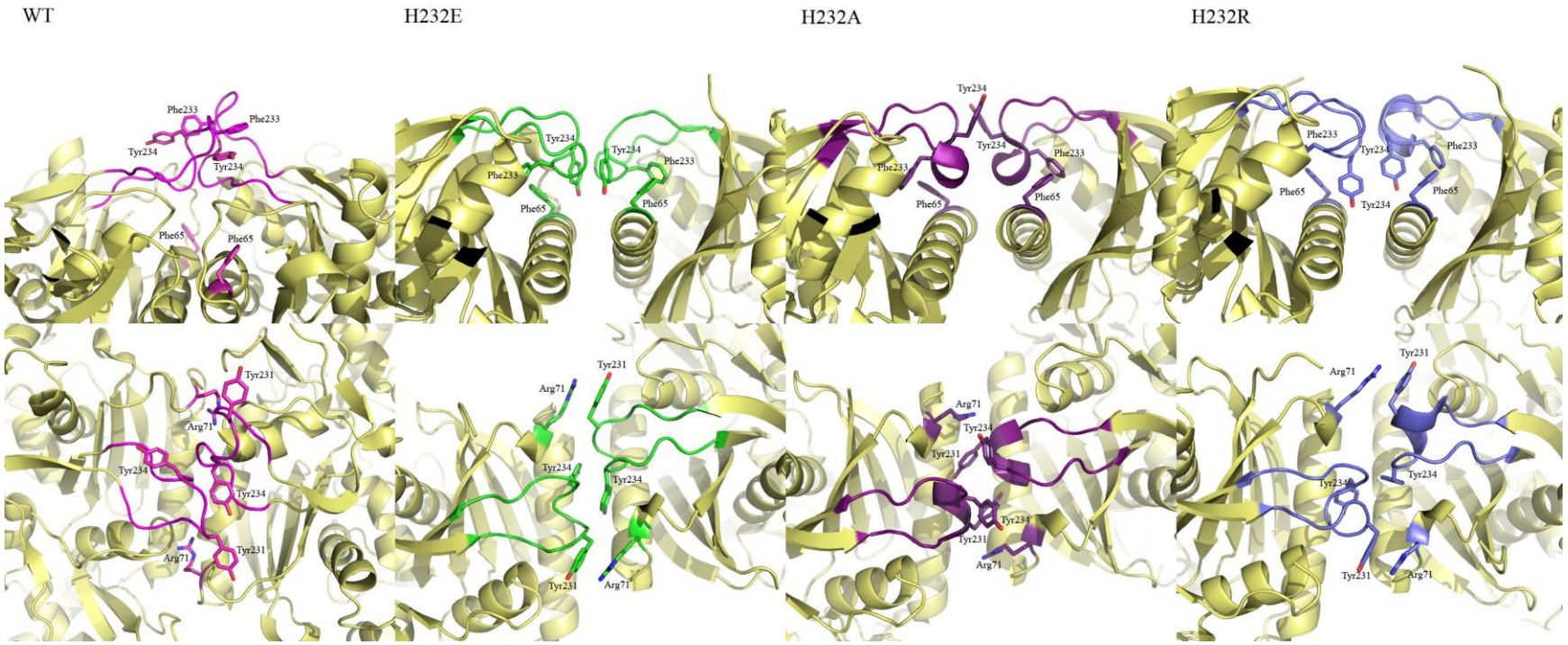

Figure 2.13 Comparison of E. casseliflavus GK loops shows the large changes in the hydrophobic residues caused by the His232Arg and His232Glu mutations

2.3.6 The relationship between active cleft volume and activation

In order to identify differences across multiple GK crystallization conditions and activation states, I used the program ^{3}V through the web interface to determine the channel size of the activation cleft in GK models [115]. Smaller channels would theoretically indicate that the model showed a more active GK or a model that was closer to the transition state than models with a larger channel. ^{3}V was used to determine the channel size of the top two or three channels in the GK models, denoted as c1 (channel 1), c2 (channel 2) and c3 (channel 3) in Table 2.4 to 2.16. ^{3}V uses a grid system to identify the channels size. ^{3}V splits up the unit cell into 1Å cubes, and then determines how cubes are involved in the channel described by the inner and outer dimensions (3Å and 10Å in the default case). Unfortunately, using the entire deposited PDB for channel searches gives results that are not comparable from model to model (see Table 2.17). In cases where two channels are approximately the same size, ^{3}V outputs the data for both channels plus a third channel so that the user can identify which channels are relevant to analysis. Initial calculations were done with the deposited PDBs, with the hope that ^{3}V would be able to identify and output the channels in each monomer. Results varied, with some monomer channels being identified, but other cases showed the entire area between the monomers as a channel. Some monomers in the PDB are separated enough from each other so the program identifies each monomer as having its own channel. Some PDBs have monomers that are close enough so the program calculates the entire inside area between the monomers as a channel. The inconsistency creates statistics that are not comparable. In order to obtain comparable numbers and ensure that the described channel included the active site, the PDBs were split into separate monomers. The results are listed in Table 2.18. ^{3}V outputs the channel volume and surface area for the largest two channels, as well as the coordinates for the center of each calculated channel.

The program provides image files that can be loaded into a visualization program to show the channel placement in the model. The channel images show that the described channels did not always extend to the glycerol molecule, suggesting that the entire channel size was not determined. In order to determine the channel size back to the glycerol molecule, the minimum radius was decreased to 2 Å, a little bit larger than the radius of 1.5 Å that water can enter. The results are shown in Table 2.19.

RMSD analysis of the *P. aeruginosa* GK tetramer suggests that it is closest to the structure of the inactive tetramer PDB ID 1BU6. The smallest rotation is necessary to superimpose Domain II after superposition of Domain I. The rotation necessary to superimpose the ethylene glycol-bound GK suggests that the active site of the ethylene glycol-bound *P. aeruginosa* GK is more open than the glycerol-bound structure, which matches the results of the *En. casseliflavus* GK.

Table 2.17. Volume and surface area of channels in deposited PDB, calculated by ^{3}V.

	Volume[a] ($Å^3$)	Surface Area[b] ($Å^2$)	Center of Mass[c] (x, y, z)	Reduced Center of Mass[d] (x, y, z)
1BWF c1	911	577	(35.2, 14.7, 36.3)	(35.1, 15.0, 37.1)
1BWF c2	1167	886	(24.9, 47.0, 27.4)	(22.7, 46.7, 23.1)
1BOT c1	899	583	(46.9, 31.6, 118.8)	(45.0, 33.3, 117.3)
1BOT c2	336	301	(15.1, 38.5, 110.2)	(14.9, 38.1, 110.3)
1BO5 c1	696	557	(64.9, 34.6, 103.4)	(62.7, 36.9, 104.4)
1GLF c1	538	414	(39.3, 12.9, 51.7)	(40.0, 11.0, 53.0)
1GLF c2	1757	7353	(20.7, -1.1, 10.1)	(3.6, -6.8, 6.5)
1BU6 c1	14500	6642	(-24.2, 1.4, 60.5)	(-40.0, 1.5, 66.0)
1BU6 c2	523	418	(-8.1, -11.8, 27.1)	(-9.0, -10.5, 28.0)
1GLA c1	2347	1112	(-34.9, 62.2, -7.7)	(-35.4, 62.3, -7.9)
1GLA c2	2347	1112	(47.9, 62.2, -5.3)	(48.4, 62.3, -5.1)
1GLA c3	73664	23705	(1.0, 62.5, -1.0)	(-26.0, 64.9, -6.6)
1GLB c1	1975	948	(-37.1, 62.5, -5.1)	(-39.1, 64.2, -6.0)
1GLB c2	1975	948	(44.1, 62.5, -1.9)	(41.9, 60.6, -1.0)
1GLB c3	73722	24253	(3.0, 62.4, -3.0)	(-25.0, 65.0, -8.5)
1GLJ c1	870	555	(34.9, 14.2, 36.2)	(34.0, 11.0, 33.0)
1GLJ c2	1032	755	(36.4, 18.6, 8.6)	(36.4, 17.9, 7.6)
1GLL c1	875	555	(34.8, 14.8, 36.5)	(34.0, 12.0, 35.0)
1GLL c2	926	657	(34.4, 18.5, 8.3)	(34.9, 18.1, 7.9)
1XUP c1	2839	1337	(-6.9, -14.6, -16.7)	(-7.0, -15.9, -17.0)
1XUP c2	662	435	(-14.0, -33.3, -42.5)	(-14.2, -33.6, -43.1)
1R59 c1	4317	2051	(22.2, 63.8, 97.4)	(21.1, 64.7, 100.1)
1R59 c2	718	461	(32.7, 97.0, 72.8)	(31.6, 98.3, 72.8)
3H3N c1	747	572	(-14.9, 40.3, 32.9)	(-15.0, 40.5, 32.0)
3H3N c2	4915	2607	(-1.9, 19.6, 14.6)	(-8.9, 17.7, 4.0)
3H3O c1	609	451	(-44.5, 23.4, 54.5)	(-44.7, 23.4, 54.7)
3H3O c2	5369	2652	(-11.5, 14.5, 49.2)	(-14.0, 14.0, 48.0)
3H45 c1	577	424	(-43.5, 2.7, 52.0)	(-45.0, 3.0, 53.0)
3H45 c2	5503	2835	(-13.6, -6.2, 49.8)	(-12.4, -6.5, 49.8)
3H46 c1	5259	2602	(-1.1, 20.5, 15.0)	-3.0, 23.0, 15.0)
PAGK EG	2014	1161	(-2.0, 11.5, 24.3)	(-1.7, 9.3, 27.0)
PAGK GOL c1	541	415	(-6.5, -9.4, 18.2)	(-6.5, -9.4, 18.2)
PAGK GOL c2	427	306	(7.1, -15.5, 10.4)	(8.0, -16.0, 11.0)

[a] The PDB ID delineates that PDB used to calculate the two largest channels. c1 and c2 describe the first and second channel, respectively, calculated.

[b] Volume of the channel is calculated by determining the number of $1Å^3$ cubes that are occupied by the channel.

[c] Surface area of the channel is calculated by counting the number of $1Å^3$ cubes on the surface of the channel that are occupied, multiplied by a weighting factor (see [115]).
[d] The center of mass is the (x, y, z) position of the average point of all of the $1Å^3$ cubes that comprise the channel. Because it is an average position, the point may reside outside of the calculated channel if the channel is sufficiently curved. Positions are in Å coordinates, not fractional coordinates.
[e] The reduced center of mass is the (x, y, z) position that results when the $1Å^3$ cubes are iteratively removed from the surface of the channel. The remaining cube is the position of the reduced center of mass. Positions are in Å coordinates, not fractional coordinates.

Table 2.18. Volume and surface area of channels in GK monomers determined via ^{3}V with an inner radius of 2Å

	Volume[a] (Å^3)	Surface Area[b] (Å^2)	Center of Mass[c] (x, y, z)	Reduced Center of Mass[d] (x, y, z)
1BWF mon1 c1	785	502	(52.2, 52.2, 13.5)	(53.8, 50.7, 14.3)
1BWF mon1 c1	911	577	(35.2, 14.7, 36.3)	(35.1, 15.0, 37.1)
1BWF mone c2	303	239	(42.4, 18.7, 15.4)	(42.7, 18.5, 15.2)
1GLJ mon1 c1	198	188	(32.1, 44.6, 23.0)	(32.0, 45.2, 22.7)
1GLJ mon1 c2	186	180	(20.6, 51.1, 20.6)	(20.4, 50.9, 21.1)
1GLJ mon1 c3	766	498	(52.1, 52.4, 13.3)	(53.6, 51.2, 14.3)
1GLJ mon2 c1	870	555	(34.9, 14.2, 36.2)	(34.0, 11.0, 33.0)
1GLJ mon2 c2	370	284	(41.8, 18.9, 16.2)	(42.0, 18.0, 16.0)
1GLL mon1 c1	210	202	(32.5, 43.9, 21.7)	(31.0, 43.0, 20.0)
1GLL mon1 c2	790	519	(51.3, 52.2, 14.4)	(52.6, 50.6, 15.4)
1GLL mon2 c1	875	555	(34.8, 14.8, 36.5)	(34.0, 12.0, 35.0)
1GLL mon2 c2	393	280	(41.3, 18.5, 15.8)	(42.0, 18.0, 16.0)
1BOT mon1	940	601	(27.8, 45.8, 86.4)	(26.0, 46.0, 82.0)
1BOT mon2 c1	163	162	(67.9, 61.7, 122.7)	(67.6, 61.5, 122.9)
1BOT mon2 c2	756	494	(64.7, 36.6, 103.2)	(62.0, 38.0, 102.0)
1BO5 mon1 c1	207	192	(22.9, 39.2, 100.4)	(22.6, 39.5, 100.2)
1BO5 mon1 c2	577	411	(30.4, 45.1, 87.3)	(29.0, 46.0, 85.0)
1BO5 mon2 c1	186	193	(62.9, 55.5, 111.6)	(63.0, 54.0, 111.0)
1BO5 mon2 c2	498	358	(64.7, 36.2, 102.5)	(64.6, 36.9, 102.4)
1GLF mon1 c1	503	359	(13.4, 9.0, 34.1)	(13.0, 8.0, 34.0)
1GLF mon1 c2	215	207	(0.1, -6.9, 29.5)	(-1.0, -6.0, 27.0)
1GLF mon2 c1	238	215	(-5.4, 4.0, 9.3)	(-6.3, 3.8, 9.7)
1GLF mon2 c2	493	398	(5.7, -11.8, -4.0)	(6.2, -11.0, -4.1)
1GLF mon3 c1	203	205	(33.0, -6.6, -0.8)	(32.4, -6.7, -1.7)
1GLF mon3 c2	558	397	(18.8, 8.5, -5.3)	(19.0, 7.3, -5.7)
1GLF mon4 c1	163	155	(49.3, -26.9, 29.2)	(49.4, -26.9, 29.1)
1GLF mon4 c2	466	349	(29.6, -9.9, 29.4)	(30.0, -9.3, 28.3)
1BU6 mon1	559	385	(-25.9, -8.1, 85.8)	(-26.5, -8.0, 86.6)
1BU6 mon2	516	369	(-16.9, 10.9, 51.1)	(-16.4, 9.8, 46.2)
1BU6 mon3 c1	219	180	(-56.8, -25.9, 46.4)	(-57.2, -25.6, 46.2)
1BU6 mon3 c2	531	392	(-32.6, -7.9, 46.0)	(-37.0, -10.0, 48.0)
1BU6 mon4 c1	491	343	(-44.6, 9.7, 82.0)	(-44.5, 9.4, 82.2)
1BU6 mon4 c2	213	184	(-63.0, -13.0, 72.4)	(-63.1, -13.1, 72.2)
1XUP mon1	566	384	(-12.2, -22.1, 17.4)	(-12.0, -22.0, 16.5)
1XUP mon2 c1	264	267	(-14.7, -35.4, -33.2)	(-16.9, -36.6, -35.6)
1XUP mon2 c2	662	435	(-14.0, -33.3, -42.5)	(-14.2, -33.6, -

				43.1)
1R59 mon1 c1	221	188	(7.1, 62.5, 102.0)	(7.3, 62.4, 101.8)
1R59 mon1 c2	163	161	(26.5, 66.7, 98.8)	(25.0, 65.0, 99.0)
1R59 mon1 c3	423	311	(35.9, 52.7, 93.5)	(35.0, 52.5, 94.0)
1R59 mon2 c1	718	461	(32.7, 97.0, 72.8)	(31.6, 98.3, 72.8)
1R59 mon2 c2	658	403	(14.6, 72.8, 68.3)	(14.3, 74.1, 67.5)
3H3N mon1 c1	628	423	(8.3, 21.2, 22.9)	(7.0, 20.5, 22.9)
3H3N mon1 c2	591	428	(29.0, 16.2, 2.5)	(29.0, 16.7, 2.5)
3H3N mon2 c1	626	438	(-33.4, 8.0, 30.2)	(-33.8, 8.8, 29.5)
3H3N mon2 c2	585	384	(-10.8, 17.6, 4.9)	(-11.8, 17.9, 4.9)
3H3O mon1 c1	456	351	(10.0, 10.1, 56.1)	(10.0, 10.0, 56.0)
3H3O mon1 c2	854	540	(4.7, 23.9, 43.1)	(5.5, 22.5, 43.1)
3H3O mon2 c1	609	451	(-44.5, 23.4, 54.5)	(-44.7, 23.4, 54.7)
3H3O mon2 c2	690	459	(-17.3, 5.5, 42.8)	(-16.5, 6.8, 42.6)
3H3O mon3 c1	697	474	(49.5, 5.9, -18.9)	(48.4, 7.4, -18.7)
3H3O mon3 c2	443	338	(68.1, 30.0, -26.7)	(68.7, 30.1, -26.9)
3H3O mon4	590	382	(23.9, 17.8, -20.0)	(23.6, 17.1, -19.7)
3H45 mon1 c1	417	310	(8.5, -10.9, 55.2)	(8.1, -10.8, 55.1)
3H45 mon1 c2	967	615	(1.9, 2.9, 44.2)	(0.0, 6.0, 42.0)
3H45 mon2 c1	577	424	(-43.5, 2.7, 52.0)	(-45.0, 3.0, 52.0)
3H45 mon2 c2	692	456	(-18.6, -15.2, 40.8)	(-16.6, -14.4, 41.6)
3H45 mon3 c1	650	407	(49.3, -13.5, -17.9.)	(47.3, -13.1, -18.6)
3H45 mon3 c2	626	440	(69.2, 8.9, -26.8)	(70.0, 9.0, -27.1)
3H45 mon4 c1	613	391	(22.4, -3.8, -18.8)	(20.5, -4.2, -17.8)
3H45 mon4 c2	542	376	(9.1, -25.2, -31.5)	(10.4, -25.5, -32.1)
3H46 mon1 c1	509	347	(12.8, 19.5, 19.7)	(11.9, 19.0, 20.0)
3H46 mon1 c2	779	527	(30.2, 11.3, 0.6)	(30.0, 12.2, 0.4)
3H46 mon2 c1	638	431	(-29.6, 12.1, 29.3)	(-29.5, 13.7, 29.7)
3H46 mon2 c2	671	492	(-8.3, 21.2, 3.7)	(-7.0, 19.4, 3.9)
1GLA c1	744	508	(18.1, 56.1, 12.1)	(16.0, 58.0, 11.0)
1GLA c2	181	179	(10.4, 44.7, -2.9)	(10.0, 43.0, -3.0)
1GLA c3	221	186	(16.7, 33.9, -8.5)	(16.7, 33.7, -8.4)
1GLB c1	628	433	(17.0, 57.4, 13.5)	(15.0, 57.5, 13.0)
1GLB c2	205	201	(8.8, 44.7, -1.7)	(8.7, 45.6, -0.8)
1GLB c3	183	159	(18.1, 34.4, -9.1)	(18.3, 34.0, -9.4)

[a] The PDB ID delineates that PDB used to calculate the two largest channels. c1 and c2 describe the first and second channel, respectively, calculated.

[b] Volume of the channel is calculated by determining the number of $1Å^3$ cubes that are occupied by the channel.

[c] Surface area of the channel is calculated by counting the number of $1Å^3$ cubes on the surface of the channel that are occupied, multiplied by a weighting factor (see [115]).

[d] The center of mass is the (x, y, z) position of the average point of all of the $1Å^3$ cubes that comprise the channel. Because it is an average position, the point may reside outside of the calculated channel if the channel is sufficiently curved. Positions are in Å coordinates, not fractional coordinates.

[e] The reduced center of mass is the (x, y, z) position that results when the $1Å^3$ cubes are iteratively removed from the surface of the channel. The remaining cube is the position of the reduced center of mass. Positions are in Å coordinates, not fractional coordinates

Table 2.19. Volume and surface area of channels in GK monomers, determined via ^{3}V with an inner radius of 3Å

PDB	Volume[a] ($Å^3$)	Surface Area[b] ($Å^2$)	Center of Mass[c] (x, y, z)	Reduced Center of Mass[d] (x, y, z)
1BOT mon1 c1	381	339	(24.2, 38.5, 97.8)	(19.0, 41.0, 100.0)
1BOT mon1 c2	1208	741	(28.8, 45.7, 85.5)	(27.0, 46.0, 81.0)
1BOT mon2 c1	527	480	(66.8, 56.0, 106.4)	(62.5, 57.0, 105.0)
1BOT mon2 c2	1343	934	(67.9, 38.0, 102.6)	(65.6, 37.4, 102.4)
1BO5 mon1 c1	429	378	(23.4, 37.9, 99.9)	(22.6, 39.9, 99.4)
1BO5 mon1 c2	942	622	(26.4, 45.2, 88.5)	(26.7, 45.7, 88.2)
1BO5 mon2 c1	541	486	(66.9, 55.7, 107.4)	(68.8, 52.7, 107.8)
1BO5 mon2 c2	1037	698	(65.3, 36.5, 105.4)	(62.0, 38.0, 103.0)
1BU6 mon1 c1	1088	685	(-24.8, -8.6, 87.8)	(-28.0, -8.0, 87.5)
1BU6 mon1 c2	423	395	(-11.7, 6.1, 81.1)	(-13.0, 6.0, 83.5)
1BU6 mon2 c1	985	652	(-14.6, 11.2, 50.1)	(-13.4, 10.0, 49.4)
1BU6 mon2 c2	388	345	(-2.8, 29.9, 49.8)	(-8.0, 31.0, 46.5)
1BU6 mon3 c1	355	290	(-57.1, -26.2, 47.6)	(-59.0, -28.0, 50.0)
1BU6 mon3 c2	806	572	(-31.2, -8.7, 43.9)	(-30.8, -7.0, 42.2)
1BU6 mon4 c1	919	577	(-45.9, 10.5, 82.5)	(-46.4, 9.2, 83.6)
1BU6 mon4 c2	388	345	(-48.0, -3.9, 67.2)	(-50.0, -4.0, 65.0)
1BWF mon1 c1	482	416	(32.0, 45.3, 24.2)	(30.0, 38.0, 23.0)
1BWF mon1 c2	1085	651	(49.5, 52.3, 15.1)	(51.7, 50.7, 15.9)
1BWF mon2 c1	1910	1262	(50.3, 16.0, 16.5)	(41.0, 19.0, 18.0)
1BWF mon2 c2	476	445	(28.6, 12.8, 9.5)	(26.5, 13.8, 6.2)
1GLF mon1 c1	704	445	(13.0, 9.0, 34.7)	(12.8, 7.2, 33.6)
1GLF mon1 c2	436	399	(3.0, -5.8, 26.6)	(-2.0, -3.0, 24.0)
1GLF mon2 c1	409	350	(-6.3, 3.2, 10.0)	(-10.0, 0.2, 11.8)
1GLF mon2 c2	809	546	(5.4, -12.3, -4.8)	(6.5, -10.0, -4.5)
1GLF mon3 c1	411	677	(33.3, -5.8, -0.6)	(31.9, -6.1, -1.8)
1GLF mon3 c2	804	528	(18.5, 9.0, -4.8)	(18.6, 7.3, -4.2)
1GLF mon4 c1	957	682	(34.1, -10.6, 27.0)	(32.0, -8.0, 28.0)
1GLF mon4 c2	452	377	(43.3, 3.6, 14.6)	(41.0, 4.0, 16.0)
1GLJ mon1 c1	506	443	(29.3, 45.3, 26.9)	(28.0, 43.0, 23.0)
1GLJ mon1 c2	1074	666	(51.4, 52.0, 13.1)	(53.6, 51.0, 14.0)
1GLJ mon2 c1	220	212	(25.6, 37.9, 36.1)	(25.3, 37.6, 35.3)
1GLJ mon2 c2	2430	1688	(49.5, 15.4, 11.6)	(45.5, 19.0, 14.0)
1GLL mon1 c1	395	358	(30.2, 45.1, 24.8)	(29.0, 43.0, 22.5)
1GLL mon1 c2	1095	664	(50.6, 51.9, 14.2)	(48.0, 56.0, 11.0)
1GLL mon2 c1	1187	713	(35.7, 15.0, 34.9)	(34.5, 11.0, 32.0)
1GLL mon2 c2	1359	1054	(39.3, 15.9, 11.9)	(27.0, 13.0, 4.0)
1R59 mon1 c1	673	431	(31.5, 51.6, 97.1)	(30.0, 51.5, 99.0)

1R59 mon1 c2	815	662	(9.4, 60.3, 94.5)	(13.0, 62.0, 96.0)
1R59 mon2 c1	1490	910	(36.0, 96.4, 68.0)	(37.4, 98.4, 67.0)
1R59 mon2 c2	3122	2021	(22.3, 72.6, 68.2)	(19.3, 74.1, 60.9)
1XUP mon1 c1	677	450	(-11.3, -21.8, 16.7)	(-11.0, -22.0, 15.0)
1XUP mon1 c2	571	441	(-16.9, 6.6, 3.5)	(-21.0, 3.5, 3.0)
1XUP mon2 c1	592	449	(-6.7, -15.6, -28.5)	(-6.9, -15.1, -29.0)
1XUP mon2 c2	817	557	(-11.7, -32.9, -44.3)	(-12.3, -33.6, -45.9)
3H3N mon1 c1	830	575	(29.7, 16.7, 3.2)	(29.0, 17.1, 2.5)
3H3N mon1 c2	1268	812	(9.2, 22.3, 23.9)	(6.7, 20.7, 23.5)
3H3N mon2 c1	993	716	(-38.8, 8.7, 33.6)	(-38.4, 9.3, 34.3)
3H3N mon2 c2	1190	738	(-13.6, 18.6, 4.4)	(-12.1, 17.7, 5.4)
3H3O mon1 c1	1105	910	(9.0, 9.3, 59.7)	(9.0, 11.0, 56.0)
3H3O mon1 c2	1299	804	(5.5, 24.2, 43.2)	(6.4, 22.6, 42.6)
3H3O mon2 c1	1410	1039	(-44.2, 20.2, 60.1)	(-46.0, 24.0, 57.0)
3H3O mon2 c2	1233	769	(-23.1, 5.1, 46.3)	(-21.5, 6.7, 47.5)
3H3O mon3 c1	1126	706	(51.1, 6.0, -18.8)	(49.3, 7.1, -19.6)
3H3O mon3 c2	1430	1121	(56.3, 24.6, -20.0)	(55.5, 30.0, -14.0)
3H3O mon4 c1	1200	768	(22.1, 18.6, -18.0)	(20.6, 16.9, -17.1)
3H3O mon4 c2	1981	1518	(-1.3, 0.6, -18.4)	(-2.0, -4.9, -19.3)
3H45 mon1 c1	685	517	(9.7, -10.3, 54.6)	(9.0, -11.0, 55.0)
3H45 mon1 c2	1363	822	(3.3, 3.3, 43.9)	(0.0, 2.0, 47.0)
3H45 mon2 c1	1412	1059	(-44.3, -0.6, 58.3)	(-46.0, 3.0, 54.0)
3H45 mon2 c2	1149	695	(-20.7, -15.9, 42.5)	(-25.0, -15.0, 41.0)
3H45 mon3 c1	1254	758	(51.5, -15.0, -19.9)	(49.6, -13.0, -20.0)
3H45 mon3 c2	1477	1092	(56.6, 4.3, -19.8)	(57.9, 9.0, -14.9)
3H45 mon4 c1	1212	737	(22.7, -2.6, -19.2)	(20.6, -4.4, -18.4)
3H45 mon4 c2	1514	1107	(6.0, -21.1, -24.8)	(7.3, -25.4, -28.7)
3H46 mon1 c1	1043	680	(12.5, 20.8, 21.9)	(10.9, 18.8, 21.2)
3H46 mon1 c2	1691	1166	(22.8, 11.1, 7.8)	(19.1, 12.4, 11.1)
3H46 mon2 c1	1602	1178	(-31.2., 12.4, 30.0)	(-29.0, 12.0, 27.0)
3H46 mon2 c2	1068	724	(-10.5, 21.2, 4.8)	(-9.1, 19.4, 5.6))
1GLA c1	1026	677	(15.7, 55.4, 15.3)	(13.8, 56.4, 14.1)
1GLA c2	362	318	(16.9, 33.8, -7.3)	(17.1, 33.8, -8.1)
1GLB c1	867	566	(16.1, 56.6, 14.4)	(14.0, 54.0, 13.0)
1GLB c2	363	336	(17.2, 34.2, -6.5)	(16.0, 33.0, -2.0)
PAGK EG c1	3149	1987	(-5.4, 15.1, 25.7)	(-4.6, 9.6, 29.8)
PAGK EG c2	196	205	(6.8, 29.7, -0.4)	(6.0, 30.0, -3.0)
PAGK GOL c1	175	148	(-5.4, -9.2, 5.5)	(-5.3, -9.1, 5.5)
PAGK GOL c2	1748	1373	(-3.2, -13.4, 15.9)	(-4.3, -14.3, 22.0)

[a] The PDB ID delineates that PDB used to calculate the two largest channels. c1 and c2 describe the first and second channel, respectively, calculated.

[b] Volume of the channel is calculated by determining the number of $1Å^3$ cubes that are occupied by the channel.

[c] Surface area of the channel is calculated by counting the number of $1Å^3$ cubes on the surface of the channel that are occupied, multiplied by a weighting factor (see [115]).

[d] The center of mass is the (x, y, z) position of the average point of all of the $1Å^3$ cubes that comprise the channel. Because it is an average position, the point may reside outside of the calculated channel if the channel is sufficiently curved. Positions are in Å coordinates, not fractional coordinates.

[e] The reduced center of mass is the (x, y, z) position that results when the $1Å^3$ cubes are iteratively removed from the surface of the channel. The remaining cube is the position of the reduced center of mass. Positions are in Å coordinates, not fractional coordinates.

2.3.7 Interaction between *En. casseliflavus* GK and *En. faecalis* HPr and effects on crystallization

Crystals that were obtained from the *En. casseliflavus* GK-*En. faecalis* HPr cocrystallization only contained GK. However, unlike the previous WT GK structure, the activation loop is in the down position close to domain I. Having the loop close to the domain shows that the HPr that is present affects the GK, priming it for catalysis even when GK is not phosphorylated. Initial refinement of these crystals shows electron density is present for the GK, but electron density for HPr is not easily visible at sigma levels of 1.0.

Initially, *En. casseliflavus* GK and HPr were combined in the same tube before crystallization screening. Because those trials did not lead to a cocrystal, the association and dissociation constants of GK and HPr in the presence of various glycerol concentrations were tested by surface plasmon resonance (SPR, see Table 2.20). Previous SPR experiments by Dr. Ruth L. Saxl and Dr. Haibin Shi were done without glycerol, but crystallization trials were conducted in the presence of 10% glycerol. However, glycerol affects the SPR results such that the binding appears to be negative under all conditions except those with high HPr concentrations. A second experiment of SPR was done to account for the extra binding to the reference channel of the chip. The binding was causing negative readings of HPr binding to GK at high concentrations of glycerol. By using a different reference channel that already contained protein that the HPr would not bind to, the binding of HPr to GK at high concentrations of glycerol could be determined. The reference channel contained 1260 RU of full-length HIV

capsid protein. Using the reference channel that already contained protein, a positive RU was obtained during HPr binding with high glycerol concentrations, giving binding and dissociation constants (see Table 2.20).

Table 2.20 Association and Dissociation of En. casseliflavus GK and En. faecalis HPr using SPR

	k_a (M)	k_d (M)	K_A (M)	K_D (M)
No glycerol	$5.8x10^{2}$	$3.3x10^{-3}$	$1.8x10^{5}$	$5.7x10^{-6}$
1% glycerol	370	$7.26x10^{-3}$	$5.09x10^{4}$	$1.96x10^{-5}$
5% glycerol	24	0.01	$2.3x10^{3}$	$4.3x10^{-4}$
5% glycerol*	354	0.0186	$1.9x10^{4}$	$5.26x10^{-5}$
10% glycerol	$5.2x10^{2}$	$7.9x10^{-5}$	$6.6x10^{6}$	$1.5x10^{-7}$
10% glycerol*	58.4	$3.24x10^{-3}$	$1.8x10^{4}$	$5.56x10^{-5}$
10% glycerol**	$5.71x10^{3}$	$7.71x10^{-3}$	$7.41x10^{5}$	$1.35x10^{-6}$

*Constants were determined from a separate run using the three highest concentrations of HPr. This contained only results with positive values of GK-reference.
** Constants determined by using a channel containing immobilized protein (full-length HIV capsid) as the reference. This eliminated the negative responses seen in the previous tests.

The presence of HPr affects the position of the GK activation loop in the crystals. Although HPr interacts and binds to GK at micromolar concentrations, it is not visible in the crystals. The WT *En. casseliflavus* GK has loop contacts that are stabilizing the interactions between the activation loop and Domain I, including the stacking interactions between the aromatic amino acids as is seen in the mutants. This suggests that HPr changes the stacking interactions in the OX activation loop interface, allowing the interacting activation loops to dissociate from each other and create new stacking interactions with Domain I.

2.3.8 Interactions between GK and GlpD or GlpO and crystallization trials

Crystallization screens contained *P. aeruginosa* GlpD and *P. aeruginosa* GK in a 1:1 molar ratio. Crystals that were obtained from initial screens contained *P. aeruginosa* GK only, showing that the interaction was not stable enough under the crystallization conditions. When using a His-

tagged GlpD and a non-His tagged GK, the *P. aeruginosa* GlpD does not associate strongly with the *P. aeruginosa* GK, even in low salt. The flow through and wash of copurification trials contained a significant amount of GK that is not seen in the eluted fractions. Most of the protein in the eluted fractions are GlpD, which makes sense because the GK does not have a His tag, so there should not be spurious His-tag association of GK. Consultation with Dr. Haibin Shi, who ran the initial SPR experiments showing an interaction between the GK and GlpD in both *E. coli* and *P. aeruginosa*, and with Dr. Xinmin Li and Dr. Shoucheng Du, who are working on the *E. coli* GK-GlpD cocrystals, showed that the *E. coli* GK does not associate with the *E. coli* GlpD in the presence of Tris buffer. Crystallization in HEPES pH 7.0 provided 19 hits from four different crystallization screens. Many of these crystals were in the hits in the first round and the new hits contain crystals with the same morphology as the previously determined GK.

Buffer conditions under which the GlpD and GK will interact via a pulldown assay were tested to determine buffer dependence of the GlpD-GK interaction. If a specific buffer prevents the interaction between the two proteins, the proteins would be more likely to crystallize as a single protein and not as a complex. Identifying conditions of association increases the possibility of cocrystals and indentifies possible screening conditions to avoid. The GlpD contains a His tag, and we have constructs of GK either with or without His tag. Testing CHES pH 8.8 and CHES pH 9.0 gave inconclusive results on a native gel.

GlpO was also tested in a pull-down assay with *P. aeruginosa* GK because GlpD and GlpO both perform a similar function converting G3P into DHAP. This bypasses multiple difficulties with the GK-GlpD interaction pulldown: 1) the known transient interaction between the affinity columns and the GK, 2) the instability of the GlpD, 3) the expense of the β-

octylglucoside used in the GlpD protein preparation and 4) the necessity of purifying GlpD, regenerating the beads, and running the pulldown during the same day.

Streptococcus pyogenes (*S. pyogenes*) GlpO, purified from a commercially purchased source by Dr. Unmesh Chinte, was tested with a cobalt resin for the pulldown with *P. aeruginosa* GK containing a His tag. Trichloroacetic acid (TCA) precipitation and SDS-PAGE gels were run to determine which proteins were present. The TCA precipitation concentrates the proteins so that small amounts can be detected on a gel without using large amounts of sample. HEPES pH 6.5 and MES pH 6.5 both show interaction between the *P. aeruginosa* GK and *S. pyogenes* GlpO.

2.4 DISCUSSION

All of the GK structures show a tetrameric orientation, like previous *E. coli* structures [53, 54, 57, 58, 113]. Unlike the *E. coli* structures, the *En. casseliflavus* GK tetramer shifts to bring the bottom of Domain II of the O and X monomers closer together during activation. This shift is reflected in the entire tetramer, initiated by the loop structures and the movement of Domain I at the OX interface. All of the mutants show the same tetrameric conformation, even with multiple space groups represented, showing that crystallization contacts are not a determining factor in the tetrameric orientation. With the assumption that the *P. aeruginosa* GK is controlled similarly to the *E. coli* GK, the tetramers can be examined to determine whether or not this is an active tetramer. The tight association between the O and X monomers and comparison of the orientation of the *P. aeruginosa* and *E. coli* monomers suggest that the *P. aeruginosa* GK is in the inactive conformation. The prevalence of the inactive conformation, seen in multiple

crystallization conditions, and the stability of the crystals suggests that trapping the *P. aeruginosa* GK in the inactive conformation by stabilizing the OX dimer will be a good approach to GK inhibition.

2.4.1 Activation caused by activation loop stabilization

Activation of the Gram positive GK has been a question since mutations revealed activation linked to the activation loop [37, 59]. Previous mutations in *Bacillus subtilis* GK were analyzed with the knowledge of the *Escherichia coli* GK structure [97], but without the knowledge of the two type of tetramers that were shown to affect activity in Gram negative bacteria [57, 114] or that the activation is regulated via the activation loop rather than the interaction between GK and $EIIA^{Glc}$ [54]. The *En. casseliflavus* structures shed light on the method of activation in the *B. subtilis* mutants. Whejte et al. had described mutants that were contained in the activation loop [97]. According to the *E. coli* models ar the time, they had proposed that the mutants either stabilize or destabilize the interaction between the O and X monomers. However, we can now identify that the mutants would have affected the position and stability of the activation loop, and therefore the activity of the *B. subtilis* GK.

2.4.2 Glycerol placement and activation

The glycerol-bound models have a more closed catalytic cleft than the ethylene glycol models, suggesting that interactions specific for glycerol affect the catalysis and orientation of the active glycerol kinase. The ethylene glycol models do not contain consistent interactions with Glu85 or Asp246, and hydrogen bonding distances between the ethylene glycol and Tyr136 are larger than

in the glycerol-bound models. Larger distances are also seen in the interactions that are present with Asp246 and Glu85. The differences in the hydrogen bonding distances are indicative of weaker interactions. Weaker interactions are going to cause a smaller closure between Domain I and Domain II, preventing correct alignment of the active site constituents.

The *P. aeruginosa* GK activates in part by the addition of ATP. In these structures lacking ATP a larger region from residues 315-340 are difficult to trace. The absence of ATP demonstrates that part of the activation control in the *P. aeruginosa* GK comes from the disorder to order transition of this region, as in the *E. coli* structures [57]. We are currently attempting to cocrystallize the *P. aeruginosa* GK with ATPγS in order to confirm this hypothesis. The electron density is low and the B values are higher than the structural average, suggesting a high amount of movement in this region. In the *E. coli* models, the corresponding region of GK is structured when in the presence of ATP. The ATP binding site is composed of those residues, and superposition of the *E. coli* GK and the *P. aeruginosa* GK shows the ATP from the *E. coli* models would fit in the high movement region of the *P. aeruginosa* GK.

While ATP is necessary for the activity of *En. casseliflavus* and *P. aeruginosa* GK, the differences in the activation regulation demonstrate the necessary differences in the physiologically active dimer. In the *E. coli* GK, and presumably the *P. aeruginosa* GK, the OXYZ tetramer dissociates into an OY dimer, which is the active form. Mutations in OX dimer interface that disrupt the OX interactions showed no reduction in activity and a very different tetrameric form than previous WT models [57, 114]. However, the *En. casseliflavus* GK is physiologically active in the OX dimer. These structures demonstrate that the interactions created by the OX dimer are necessary for the domino effect caused by phosphorylating His232.

2.4.3 Controlling activation – drug targets

The difference in the GK activation mechanisms of Gram positive and Gram negative bacteria show that the interface between the O and X monomers is a prime target for control GK activation. In the Gram negative bacteria, such as the *P. aeruginosa*, stabilization of the OX dimer inhibits the activity. The fact that *P. aeruginosa* appears to consistently form the inactive tetramer in numerous crystallization conditions, creating crystals with the same space group, unit cell, and tetramer orientation, suggests that stabilization of the inactive *P. aeruginosa* GK is extremely feasible. Activation of *P. aeruginosa* GK also seems to be affected by ATP binding to Domain II, where a region of the protein is not visible in current structures. Inhibitors that would prevent ATP binding could use this disorder-to-order transition to their advantage.

In the *En. casseliflavus* GK, tactics that would dissociate the OX dimer and disrupt the OX interface would inhibit GK activation. Promoting disorder of the activation loop would also decrease GK activation. The information gleaned from the differences in the activation of the *En. casseliflavus* GK versus the *P. aeruginosa* GK suggest that focusing on the dimerization and tetramerization of GK would help identify various drug targets able to specifically inhibit GK activation.

2.4.4 Summary and Future Directions

I have shown that the mutation of the *En. casseliflavus* His232Arg creates a domino effect that closes the active cleft and leads to activation of the GK. Although the His232Glu mutation is the same charge as a phosphorylated histidine it is not long enough to reach the positively-charged lysine, preventing loop stabilization and activation. The presence of *En. faecalis* HPr is enough

to position the activation loop between the X-O dimer, but cocrystallization attempts have thus far been unsuccessful.

The presence of glycerol closes the active site more than ethylene glycol in both the *E. casseliflavus* and *P. aeruginosa* GK. These models are the first models that contain ethylene glycol, and comparison of the *P. aeruginosa* GK with *E. coli* GK shows that even in the inactive tetrameric conformation the contacts between the substrate and the protein affect the opening of the catalytic cleft.

Preliminary pull-down binding data for the *P. aeruginosa* GK and GlpD, as well as the *P. aeruginosa* GK and *S. pyogenes* GlpO, demonstrate interactions between GK and GlpD/GlpO. However, the GK crystallization is so robust that crystals large enough to diffract are *P. aeruginosa* GK only. Initial pull-downs discussed in this chapter and SPR assays by Dr. Haibin Shi suggest that HEPES buffers stabilize the interaction between *P. aeruginosa* GK and GlpD/GlpO. Further cocrystallization attempts should be undertaken utilizing this information to determine the regulation and interaction between GK and GlpD/GlpO.

3.0 OCTAMERIC HEMOGLOBIN

3.1 INTRODUCTION

Human hemoglobin is a tetrameric oxygen carrier found in blood cells. It is the component responsible for delivering oxygen to tissues, and must be able to obtain oxygen from oxygenated lungs. The monomers of adult human hemoglobin (HbA) is very similar in structure to myoglobin, the monomeric oxygen-storage protein found in muscles. With myoglobin, hemoglobin was one of the first x-ray crystallographic structures determined.

3.1.1 Hemoglobin structure organization and nomenclature

Hemoglobin is listed as an α protein, meaning that the majority of the secondary structure consists of α helices. The α and β monomers have the same number of helices, listed alphabetically from A to H [3, 4]. Table 3.1 shows the helix designation and residues.

Table 3.1 Helix designations for α and β chains adapted from references. [3, 4]

Helix designation	α helix residues	β helix residues
A	3-17	5-17
AB	18-22	18-22
B	23-35	23-34
C	36-43	35-42
CD	44-52	43-49
D	--	50-56
E	53-70	57-77
EF	71-74	78-79
F′	75-79	80-84
F	80-90	85-95
FG	91-97	96-103
G	98-112	104-117
GH	113-117	118-122
H	118-124	123-139

The porphyrin group sits in a hydrophobic pocket in each subunit, with its carboxyl tails exposed to the solvent. The large mobile tails on the same side of the hydrophobic porphyrin reinforce the correct orientation of the heme within each subunit.

3.1.2 Mutants affecting views of allostery and hemoglobin activity

HbA undergoes a series of transitions when binding ligand. HbA can bind oxygen or carbon monoxide each heme group when the heme iron is reduced to the Fe^{2+} state, and is able to bind cyanide, water, or hydroxide when oxidized in the Fe^{3+} state. When unliganded, the heme adopts a T state structure. The iron is coordinated with the four nitrogens of the porphyrin group and the NE of the proximal histidine (αHis87 and βHis92), but is slightly out of the plane of the porphyrin [2, 12]. Binding of a ligand, whether oxygen, carbon monoxide, cyanide, water, of

hydroxyl, will cause the heme group to adopt the R conformation, where the iron is in the plane of the porphyrin [2-4, 7, 12]. The T and R states also demonstrate a large change in the relative positions of the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers. Superposition of the T state $\alpha_1\beta_1$ dimer and the R state $\alpha_1\beta_1$ dimer shows that a 15° rotation of the $\alpha_2\beta_2$ dimer would be necessary to superimpose the second dimer [2-4, 7, 12]. Further crystallization demonstrated that the simple T-R dichotomy of HbA structures is incorrect, identifying an R2 and R3 state [6, 17]. The R2 state contains an $\alpha_2\beta_2$ dimer rotation in the same relative direction as the T to R rotation, but rotates the $\alpha_2\beta_2$ dimer a further 5 degrees [17]. The R3 state shows a different rotation of the $\alpha_2\beta_2$ dimer, showing that the tetrameric HbA can adopt a series of conformations [6]. A fourth structure, RR2, has an intermediate $\alpha_2\beta_2$ dimer rotation between the R and R2 structures [6]. Research has shown that each of these conformations can be stabilized by crystallization conditions [2-4, 6, 7, 10, 12, 16, 17, 116]. Solution state research by Dr. Ho and coworkers shows that HbA adopts a series of conformations across the R, RR2 and R2 states, suggested as the R ensemble or R^e state [117, 118].

Comparing the high-salt R structure to the R2 structure obtained from high molecular weight PEG, the R2 structure shows an increased rotation of dimers in comparison to the T state, but has an alternate β_2His97 conformation that allows more water in the $\alpha_1\beta_2$ interface than the R structure [17, 116]. The high salt RR2 structure, obtained from a more physiological pH of 6.4-7.1, is an intermediate structure between R and R2 [6]. NMR work on HbA and relevant mutants demonstrated that HbA adopts an ensemble of states between R and R2, and the most likely structural transitions during ligand binding are T->R->RR2->R2 [117, 118].

3.1.3 Blood substitutes

The shelf life of donated blood is approximately 4-6 weeks, and recent research suggests that degradation of components leading to adverse reactions start as soon as two weeks after donation. Additional complications include antigen reactivity and Rh factors that necessitate matching the donor blood to the blood type of the recipient. The storage conditions necessary for keeping blood also pose a challenge to places that are far from storage such as military units or emergencies [33] Multiple strategies have been used to attempt to create oxygen carriers: membrane vesicles, polymerization, designed multimers and various cross-linkage reactions [34, 119-128]. Researchers have been examining hemoglobin-based oxygen carriers (HBOC's) as a solution to the challenges posed by some of the current oxygen carriers. By creating HBOC's that can be produced when necessary, the need for blood donation is eliminated, prevents reactivity seen in normal blood donation, and reduces the requirements for storage. The Hb must have a low enough affinity for oxygen that the oxygen will be released, but not so low that the Hb will quickly oxidize to methemoglobin (metHb) or aquamethemoglobin (aquametHb). Unfortunately, the oxygen affinity is inversely correlated with the propensity of the heme group to become oxidated [129, 130].

The United States Navy is interested in designing blood substitutes that would be able to have an extended life span and increased availability [33]. Current trials of blood substitutes have been stopped because current blood substitutes are not advantageous over the current techniques [33]. They require a high concentration of blood substitute and show vasoactivity that can lead to heart attacks when used in the field [33]. Vasoactivity can be defined as either an expansion or contraction of blood vessels, which often puts more strain on the cardiovasculature. Vasoactivity problems seem to stem from the fact that unsequestered Hb

interacts with nitric oxide. Current proposals suggest adding nitroglycerin to avoid these difficulties but research is ongoing [33]. Increasing the size of hemoglobin or hemoglobin vesicles decreases the vasoactivity. Another difficulty in obtaining a hemoglobin-based blood substitute is the dissociation of tetrameric Hb at low concentrations. Octamerization of HbA could prevent dissociation, as discussed in Section 3.1.4.

Previous cross-linked hemoglobin structures used linkers to cross between the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers in the same tetramer and prevent dimer dissociation. Hemoglobin linkers were connected through either β_1Lys82-β_2Lys82, β_1Lys82-β_2Val1, α_1Val1-α_2Lys99 or α_1Lys99-α_2Lys99 [18, 131-145]. Previous structures that were cross-linked at the β subunits eliminated access to the 2,3-BPG binding pocket; the linker crossed over the binding pocket, in each case covalently bonding to at least one of the necessary lysines from the 2,3-BPG pocket [131-137, 145]. Using the α subunits to cross-link and stabilize the tetrameric form also greatly reduced or eliminated binding and the biochemical effects of 2,3-BPG [143, 144].

3.1.4 Octameric hemoglobin

One tactic for creating HBOC's is to link multiple tetramers together. At low concentrations, HbA will dissociate into αβ dimers which are cleared from the blood by interacting with haptoglobin (Hpt) [146]. Hpt binds to HbA dimers to prevent nitric oxide scavenging. Hb-bound Hpt signals to macrophages to clear the HbA from the blood. By creating mutations that stabilize a larger multimer, dissociation, Hpt scavenging, and nitric oxide should be decreased.

3.1.5 Mutant rationale

Only one example of a naturally-occurring octameric Hb has been discovered. A natural mutation in the β subunit (βGly83Cys) was discovered in 1971 [147], which was found to octamerize the hemoglobin. The octamer was presumed to occur via a disulfide bridge created between the two β monomers [119, 148]. Other naturally-occurring mutations to Cys include Hb Porto Alegre (βSer9Cys), Hb Whitmire (βGly16Cys), Hb Kent (βTrp37Cys), Hb Ilmenau (βPhe41Cys), βPhe42Cys, Hb Mississippi (βSer44Cys), Hb Arta (βPhe45Cys), Hb Colima (βSer49Cys), Hb Leeds (βGly56Cys), Hb Harrow (βPhe118Cys), Hb Montfermeil (βTyr130Cys), Hb Visayan (βGly136Cys), Hb Rainier (βTyr145Cys), Hb Teterboro (αSer3Cys), Hb Bladensburg (αTrp14Cys), Hb Nigeria (αSer81Cys), Hb Cap d'Agde (αSer131Cys), Hb Nunobiki (αArg136Cys), Hb St. Rose (αGly15Cys), Hb Lima (αGly18Cys), Hb Ramona (αTyr24Cys) and Hb Ecuador (αSer138Cys) [149]. However, only Hb Ta-Li (βGly83Cys) creates a controlled octameric system, in part because many of the Cys mutations are not at the outer surface of Hb. Dr. Chien Ho and coworkers created a plasmid with the βGly83Cys mutation, which yields octameric Hb that is refractory to dissociating into tetramers [34, 119], The presence of the stable octameric hemoglobin eliminated the problem of hemoglobin dissociation that would occur if HbA was just injected into blood vessels. A mutation from glycine to a cysteine was introduced at β83, creating disulfide bridges between Hb tetramers [119]. A secondary mutation was introduced at β41, from phenylalanine to tyrosine, which further altered the oxygen binding properties [34]. Both mutants are stable in octameric form, do not react with Hpt, and are not reduced to tetrameric hemoglobin in plasma [34, 119]. What is more interesting is that the octameric mutants also have altered biochemical properties. Both

mutants have slightly less cooperativity than HbA, as demonstrated by the Hill coefficient [34, 119]. Normal HbA has a Hill coefficient of around 2.6, where higher Hill coefficients are indicative of more cooperativity and a Hill coefficient of 1 indicates no cooperativity [150]. Early biochemical tests suggested that the mutation increased the relative oxygen affinity of the βGly83Cys mutant, so a secondary mutation was designed at the βPhe41 position to give a double mutant of βGly83Cys-Phe41Tyr. The βPhe41Tyr mutation decreased the oxygen affinity relative to HbA. A third mutant was designed to create octamers using the α subunit by mutating the comparable residue in the α subunit to Cys (αAsn78Cys). The α mutation is in the same relative placement as the βGly83Cys mutation – on helix F, position 7. I have determined the X-ray crystallographic structure of octameric hemoglobin in its oxidized water-bound aquamet state. Comparison of those structures with two separately determined structures of cyanomet octameric hemoglobin demonstrates how the octamerization affects the overall structure, and hence binding of ligands.

3.2 MATERIALS AND METHODS

3.2.1 Expression and Purification of rHb

Ho and coworkers expressed recombinant βGly83Cys, βGly83Cys-Phe41Tyr, and αAsn78Cys in an IPTG-inducible plasmid in *E. coli* as in Shen et al. [151]. In brief, cells were grown and induced with IPTG, sonicated and centrifuged. The supernatant was saturated with CO, and the nucleic acids were precipitated with addition of 10% polyethyleneimine to a final concentration of 0.5%. After centrifugation, the hemoglobin in the supernatant was purified using three

columns: Q-Sepharose fast flow with 20 mM Tris-HCl, 0.1 mM TETA, pH 7.4; Q-Sepharose fast flow with 20 mM Tris-HCl, 0.1 mM TETA pH 8.3, using a salt gradient to a final salt concentration of 160 mM NaCl; after dialysis into 10 mM sodium phosphate, the third column was a Mono S with a pH gradient from 10 mM sodium phosphate, 0.1 mM EDTA, pH 6.8 to 20 mM sodium phosphate, 0.1 mM EDTA, pH 8.3. The hemoglobin was further purified with a Sephadex 200 column in sodium phosphate buffer pH 7.0. All buffers were bubbled with CO prior to purification to prevent heme oxidation.

3.2.2 Secondary purification of oHb after freezing

βGly83Cys, βGly83Cys-Phe41Tyr and αAsn78Cys were obtained from Dr. Chien Ho and coworkers, purified as in Shen et al. [151]. Protein was stored at -80°C until use. Protein was thawed on ice and buffer exchanged using a Superdex 200 column. Previous work by Dr. Shoucheng Du had suggested that the octameric hemoglobin contained multiple forms of octamers and that the resulting heterogeneity affected the crystals. Plate-like crystals diffracted poorly (approximately 4-8 Å resolution) and the size and quality varied between preparations. βGly83Cys eluted from the column at a volume consistent with octameric hemoglobin. Protein was concentrated to 50 mg/mL using a 50 kDa cutoff Amicon centrifuge filters, centrifuging in an eppendorf 5810R for 10-20 min each time. Whenever possible, protein was used the day that it was purified and concentrated. Excess protein was flash frozen in liquid nitrogen at 100 K and stored at -80°C for future use.

7.5% polyacrylamide native gels of the Superdex eluent showed that the βGly83Cys and βGly83Cys-Phe41Tyr mutants were most stable in 10-50 mM Tris pH 7.5, and αAsn78Cys was most stable in 10-50 mM sodium phosphate pH 7.5 (see Figure 3.2 for an example of the

dissociation of αAsn78Cys). Native gels were used instead of denaturing SDS gels because the SDS gels would only give the presence of α and β monomers in the presence of SDS and BME. Identification of octamers and tetramers requires that the multimers are not dissociated in the preparation and gel running. Native gels were run at constant voltage of 100 V to prevent the gels from heating and denaturing the hemoglobin. After determining the stable buffer conditions, proteins were buffer exchanged using 4 mL Amicon filters with a 50 kDa cutoff, which did not affect the multimerization of the protein as demonstrated by native gels.

3.2.3 Crystallization

Hemoglobin protein was concentrated to 50 mg/mL prior to crystallization. 96-well crystal trays were originally set up using drops of 0.5 µL of protein and 0.5 µL of crystallant, and 80 µL of precipitant in the well. Initial crystallization hits included the Emerald Wizard II 18, consisting of 20% w/v PEG 3000, 0.1 M Tris pH 7.0, 0.2 M calcium acetate. Crystals of βGly83Cys came from drops of 20% w/v PEG 3000, 0.1 M Tris pH 10.0, 0.2 M calcium acetate, and crystals of βGly83Cys-Phe41Tyr came from drops of 20% w/v PEG 3000, 0.1 M Tris pH 9.0, 0.2 M calcium acetate (see Figure 3.3 and Figure 3.4). Early crystals were plates and difficult to cryoprotect and freeze without causing undue damage to them and obtaining poorly diffracting crystals. For further screening, large drops were produced using 50 mg/mL protein and manipulating the pH and PEG concentrations (see Table 3.2). Additive screens did not improve the crystal morphology or diffraction for the βGly83Cys or the βGly83Cys-Phe41Tyr proteins. αAsn78Cys was crystallized in 0.17 M lithium sulfate, 85 mM Tris-HCl pH 8.5, 25.5% PEG 4000, and 15% glycerol (see Figure 3.5).

Seeding trials using crushed βGly83Cys crystals did not improve the crystals. Seeds were created using a seeding kit from Hampton Research and approximately 10 βGly83Cys crystals from a 25% PEG 3350, 0.1 M Tris pH 10.0, 0.2 M sodium chloride condition into about 50 µL of crystallant. The seeds were created by centrifuging the stock solution in the 1.5 mL eppendorf tube with the plastic bead supplied by the kit. Dilutions of the seed stock varied from 5X to 50X. 1 µL of diluted seed stock would be combined with 4 µL of 50 mg/mL protein and 5 µL of crystallant. Seeds increased the number of single crystals present in the drops but did not improve the resolution or morphology of the crystals.

Crystal optimization screens from Hampton research, additive screens from commercial sources and the lab and seed beads were used to attempt to improve the crystals. Noting that these mutants initially crystallized in mother liquor that contained PEG as a precipitant, a PEG pH screen was created to optimize pH and buffer conditions. Initial results also showed that the βGly83Cys mutants crystallized readily in the presence of phosphate, but phosphate inhibited αAsn78Cys crystallization. The rationale behind the PEG screen was to test different buffers at different pH's in order to determine how much the buffer content rather than the buffer pH affected the crystals (see Table 3.2 for buffer conditions). All solutions contained 25% PEG 3350 as a precipitant, allowing the solutions to have a salt of choice added to them without significantly decreasing the concentration of PEG.

3.2.4 Creation of oHbCN mutants and crystallization

Early results demonstrated that the crystals did not contain carbon monoxide (see results); instead, the models showed an oxidized recombinant Hb (oHb). A wavelength scan from 420 nm to 750 nm suggested that the protein was already partially oxidized after thawing (see Figure

3.1 for oxidation spectra). In order to attempt to obtain reduced recombinant Hb, various reducing agents were added to the gel filtration and crystallization steps to try to keep the protein reduced. One challenge of working with the octameric hemoglobin was the assumption that the octameric hemoglobin is stabilized in the octameric state by disulfide bridges. This assumption meant that using a reducing agent that was too harsh would dissociate the octamers into tetramers.

Other attempts at obtaining rHb with ligands consisted of bubbling carbon monoxide (CO) through the mother liquor (also used as the crystallant) for 20 min prior to use. Crystals that were obtained from these conditions were oxidized.

Cyanide-bound hemoglobin was created by first completely oxidizing the hemoglobin using potassium ferricyanide in a 5:1 molar ratio to heme. The potassium ferricyanide was buffer exchanged using Amicon filters as described in 3.2.2, and the protein was put into Tirs or phosphate buffer pH 7.5 with 10 µM sodium cyanide for crystallization. Protein was concentrated to 50 mg/mL and used to set up crystallization trays as in 2.2.8.

Cyanide–bound βGly83Cys crystals were obtained from Wiz I 33 (2.0 M ammonium sulfate, 0.1 M CAPS pH 10.5, 0.2 M lithium sulfate) and Wiz I 28 (20% PEG 3000, 0.1 M HEPES pH 7.5, 0.2 M sodium chloride). Crystals were cryoprotected with 30% ethylene glycol and 22% ethylene glycol, respectively (see Figure 3.6 and Figure 3.7).

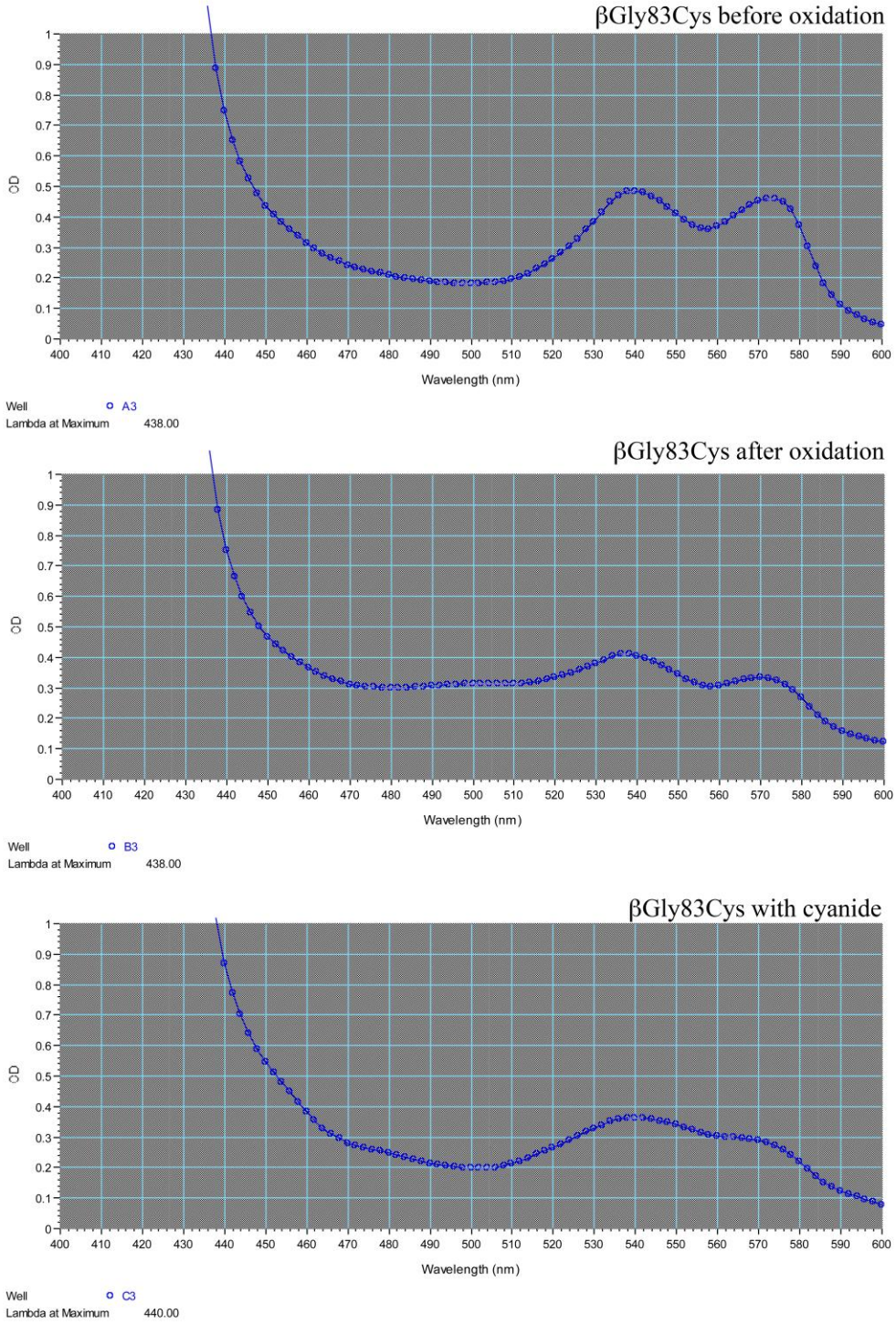

Figure 3.1 Spectra from 400 nm to 600 nm of βGly83Cys after thawing, oxidation, and treatment with cyanide.

Table 3.2 PEG pH and buffer screen

Buffer number	Precipitate	Buffer type (0.1 M)	Buffer pH
1	25% PEG 3350	Sodium citrate	4.2
2	25% PEG 3350	Sodium citrate	4.5
3	25% PEG 3350	Sodium citrate	5.0
4	25% PEG 3350	Sodium citrate	5.5
5	25% PEG 3350	Sodium acetate	4.2
6	25% PEG 3350	Sodium acetate	4.5
7	25% PEG 3350	Sodium acetate	5.0
8	25% PEG 3350	Sodium acetate	5.4
9	25% PEG 3350	Sodium acetate	5.5
10	25% PEG 3350	Sodium phosphate/citrate	4.2
11	25% PEG 3350	Sodium phosphate	6.0
12	25% PEG 3350	Sodium phosphate	6.5
13	25% PEG 3350	Sodium phosphate	7.0
14	25% PEG 3350	Sodium phosphate	7.5
15	25% PEG 3350	Sodium phosphate	8.3
16	25% PEG 3350	Bis-Tris	5.5
17	25% PEG 3350	Bis-Tris	6.0
18	25% PEG 3350	Bis-Tris	6.5
19	25% PEG 3350	Bis-Tris	7.0
20	25% PEG 3350	Bis-Tris	7.5
21	25% PEG 3350	BTP	5.5
22	25% PEG 3350	BTP	6.5
23	25% PEG 3350	BTP	7.0
24	25% PEG 3350	BTP	7.5
25	25% PEG 3350	BTP	8.0
26	25% PEG 3350	BTP	8.5
27	25% PEG 3350	BTP	9.0
28	25% PEG 3350	BTP	9.5
29	25% PEG 3350	Tris	6.5
30	25% PEG 3350	Tris	7.0
31	25% PEG 3350	Tris	7.5
32	25% PEG 3350	Tris	8.0
33	25% PEG 3350	Tris	8.2
34	25% PEG 3350	Tris	8.5
35	25% PEG 3350	Tris	9.0
36	25% PEG 3350	Tris	9.5
37	25% PEG 3350	Tris	10.0
38	25% PEG 3350	HEPES	6.0
39	25% PEG 3350	HEPES	6.5
40	25% PEG 3350	HEPES	7.0

41	25% PEG 3350	HEPES	7.5
42	25% PEG 3350	HEPES	8.0
43	25% PEG 3350	Bicine	7.5
44	25% PEG 3350	Bicine	8.0
45	25% PEG 3350	Bicine	8.5
46	25% PEG 3350	Bicine	8.9
47	25% PEG 3350	TAPS	8.0
48	25% PEG 3350	TAPS	8.5
49	25% PEG 3350	TAPS	8.8
50	25% PEG 3350	CHES	8.8
51	25% PEG 3350	CHES	9.5
52	25% PEG 3350	CHES	10.0
53	25% PEG 3350	CAPS	9.5
54	25% PEG 3350	CAPS	10.0
55	25% PEG 3350	CAPS	10.5

3.2.5 Data Collection

The βGly83Cys datasets were originally collected on the home source, a Rigaku Superbright generator. However, the cell length of the c axis was around 400Å, meaning that depending on the orientation, the reflections would be so close together that they would not be resolvable by the processing programs. Crystals were taken to Argonne National Laboratory to be exposed to the beam from SERCAT ID and BM lines, as well as the GM/CA ID line. The initial βGly83Cys-Phe41Tyr datasets contained a long axis of approximately 400 Å and radiation damage similar to the βGly83Cys datasets. Initial datasets were collected on the Raxis IV++ and HTC image plates using the Rigaku Superbright FRE generator. The βGly83Cys and βGly83Cys-Phe41Tyr datasets that were used to determine the final models were collected at wavelength 1.03317 Å (12000 keV) at GMCA beamline ID-2 at the Advanced Photon Source at Argonne National Laboratory.

The αAsn78Cys crystals were collected on both the home source and at Argonne National Laboratory APS beamline SERCAT on the BM and ID lines. αAsn78Cys data was collected using the BM beamline at APS with 1.000 Å wavelength (12398 keV). Both βGly83CysCN datasets were collected using 1.000 Å wavelength (12398 keV) at the SER-CAT ID beamline at the Advanced Photon Source at Argonne National Laboratory.

Crystals of βGly83Cys with cyanide were collected at the ID beamline at the APS SER-CAT. Datasets oC and oHG83CW were collected at 12398. eV (1.000 Å wavelength).

3.2.6 Data processing

oHb data was processed in d*trek [102] for the βGly83Cys and βGly83Cys-Phe41Tyr mutants to 2.65 Å and 2.30 Å, respectively. HKL2000 [103] was used to process the αAsn78Cys data to 2.25 Å. Data statistics are listed in Table 3.3.

Table 3.3 Data processing statistics for oHb and oHbCN data sets

	rHb (αAsn78Cys)	rHb (βGly83Cys)	rHbCN (βGly83Cys)	rHb (βGly83Cys-Phe41Tyr)	rHbCN (βGly83Cys)
Total Reflections	94,397	125,792	303,216	203,349	206,352
Unique Reflections	56,819	71,950	47,907	105,477	20,058
Resolution (Å)	50-2.25	50-2.65	50-2.5	50-2.30	50-2.93
Space group	$P2_12_12_1$	P1	$P4_12_12$	P1	$P6_522$
a (Å)	55.3	82.8	85.2	83.2	121.8
b (Å)	114.0	83.1	85.2	83.2	121.8
c (Å)	207.6	106.2	369.4	104.3	197.3
α (°)	90	115	90	93	90
β (°)	90	95	90	110	90
γ (°)	90	90	90	91	120
Completeness (%)[a]	90.4 (64.9)	98.0 (95.3)	98.6 (99.4)	93.4 (75.1)	100 (99.8)
I/σ(I)	9.46 (1.9)	9.7 (2.0)	19.2 (2.7)	10.4 (2.0)	21.3 (4.3)
R_{sym} (%)[b]	14.3 (37.9)	7.7 (37.3)	8.3 (51.8)	8.6 (41.6)	9.3 (56.4)
Redundancy	2.7 (1.7)	1.9 (1.8)	6.4 (5.2)	2.6 (2.3)	10.3 (9.9)
Mosaicity	0.372-2.936	0.512-1.649	0.331-0.543	0.533-0.777	0.300-0.787
X^2	1.044 (0.996)	1.054 (1.066)	1.016 (0.961)	1.013 (1.040)	1.005 (1.004)
Oscillation	0.5	1	0.5	1	1
Number of Images	360	180	300	240	170
Total Oscillation	180	180	150	240	170
Number of images used	198	175	170	240	84
Oscillation Used	99	175	85	240	84

[a] Numbers in parentheses are values for the highest resolution shell (out of ten shells).

[b] $R_{sym}=\Sigma_h\Sigma_i|I(h,i)-\langle I(h)\rangle|/\Sigma_h\Sigma_i I(h,i)$ where $I(h,i)$ is the intensity values of the ith measurement of h and $\langle I(h)\rangle$ is the corresponding mean values of $I(h)$ for all i measurements.

3.2.7 Molecular replacement and Model refinement

Molecular replacement was done in Phaser [152] using the protein component of a dimeric model of CO-bound hemoglobin (PDB ID 1BBB) [17], the R2 structure determined in the presence of PEG. This dimer model should be closest to the dimer model of the octameric hemoglobin because the mutant crystals also used PEG as a precipitant.

Two initial cycles of realspace refinement in Coot were followed by restrained refinement with NCS in Refmac5 [153] and simulated annealing in PHENIX [154], before addition of hydroxyl ligands. The models were refined in PHENIX [154] before addition of waters using PHENIX [154] or Coot [109]. Three cycles of water checking/ligand addition and validation using MolProbity [155] were done. After each cycle of manual adjustment, models were refined in PHENIX [154], monitoring R_{work} and R_{free} [156]. Before adding waters and ethylene glycol, R_{work}/R_{free} for the βGly83Cys, βGly83Cys-Phe41Tyr and αAsn78Cys structures was 20.2/28.2%, 23.6/25.3% and 21.7/26.4%, respectively. Waters were added to density over 2.5σ in non-protein areas 2.5-4.0Å from hydrogen-bonding atoms (nitrogen, oxygen and sulfur) using water picking in PHENIX and Coot [109]. Ethylene glycol, ions, and Tris were added to peaks with the appropriate electron density shape and sigma level. oHb structures were manually checked using validation available in Coot [109], stand-alone and PHENIX packaged MolProbity [155], and SFcheck [157] and Procheck [158] in CCP4 [106]. Statistics can be found in Table 3.4.

3.2.8 Model analysis

Distances and angles between individual atoms were determined in Coot [108, 109]. DDMP was used to determine the relative movement of some models [110]. In order to compare the models accurately to previous models and methods, a combination of RMSD and superposition was used to compare the various models. For the complete tetramers, Cα atoms after tetrameric superposition were compared using RMSD. This procedure takes the distance between each pair of Cα atoms and averages all of the distance vector lengths. Because many of the differences determined between the T state and various R state models can be determined by a rotational shift of the $\alpha_2\beta_2$ dimer relative to the $\alpha_1\beta_1$ dimer, this can also be compared using a series of superposition and RMSD calculations. In this case, lsqman, part of the dejavu package from Uppsala, was used to superimpose the $\alpha_1\beta_1$ dimer Cα atoms and calculated the RMSD [111, 112]. The RMSD for the $\alpha_2\beta_2$ dimer was then determined using the command rmsfit without a second superposition. The RMSD then gives an overview of how large the αβ dimer rotation is between the two models, but does not give the direction of rotation.

Table 3.4 Data refinement statistics of oHb mutants and oHbCN mutants

	rHb (αAsn78Cys)	rHb (βGly83Cys)	rHbCN (βGly83Cys) 1	rHb (βGly83Cys-Phe41Tyr)	rHbCN (βGly83Cys) 2
Molecular Replacement					
Rotational Z	11.0	11.5	9.8	18.0	9.8
Translational Z	16.6	100.0	31.7	100.0	41.0
Rotational Z	9.6	13.3	11.5	14.0	9.3
Translational Z	29.6	19.8	55.9	24.8	64.6
Rotational Z	9.6	10.7	8.9	13.8	
Translational Z	42.5	27.8	56.6	31.7	
Rotational Z	11.5	14.9	9.8	15.2	
Translational Z	51.6	34.8	62.5	43.3	
Rotational Z		11.5		18.3	
Translational Z		34.7		54.3	
Rotational Z		16.0		17.6	
Translational Z		47.0		57.0	
Rotational Z		16.2		23.6	
Translational Z		48.9		70.7	
Rotational Z		13.7		21.2	
Translational Z		47.4		76.1	
Clashes	0	0	6	0	1
After Refinement					
R_{work} (%)	19.02	19.95	19.56	18.20	19.94
R_{free} (%)	23.31	24.63	25.64	22.24	23.81
Ramachandran Statistics					
Bond Lengths (Å)	0.013	0.009	0.003	0.008	0.009
Bond Angles (°)	1.124	1.055	0.69	1.016	1.12
Most Favored	92.1	91.9	97.0	92.8	95.9

Additionally Allowed	7.9	8.1	2.8	7.2	3.0
Disallowed	0	0	0.2	0	1.1
Clashscore		17.46	21.15	14.80	15.89
C_β Outliers		0	0	0	0

3.3 RESULTS

3.3.1 Determining acceptable buffer conditions

Initial crystallization attempts gave small crystals that contained a highly variable unit cell in the βGly83Cys mutants, and very small, poorly-diffracting (8-10 Å) crystals of the αAsn78Cys mutant. In order to determine how buffer conditions were affecting the crystallization, the protein was run on a Superdex 200 gel filtration column. All proteins showed results indicative of two peaks, one of the major octameric component, and one of the dissociated tetrameric component. Running the fractions on a 7.5% polyacrylamide native gel confirmed the presence of both components (See Figure 3.2 for an example of the αAsn78Cys dissociation gels). Results from the Superdex 200 columns showed that the βGly83Cys and βGly83Cys-Phe41Tyr mutants were stable in 50 mM sodium phosphate pH 7.5, while the αAsn78Cys mutant was stable in 50 mM Tris-HCl pH 7.5. For future crystallization, the proteins were buffer exchanged using Amicon filtration centrifuge tubes.

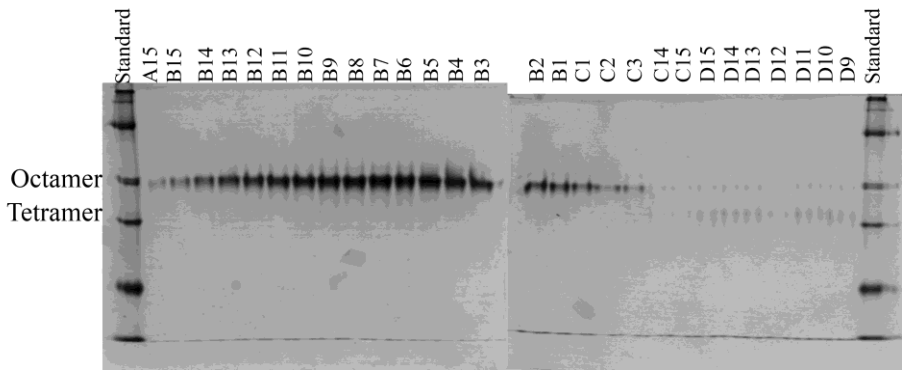

Figure 3.2 Polyacrylamide native gels (7.5%) of Superdex 200 fractions of αAsn78Cys. Fractions show that one peak consists of stable octameric αAsn78Cys, while the other peak consists of tetrameric αAsn78Cys. βGly83Cys and βGly83Cys-Phe41Tyr mutants show the same feature.

3.3.2 Crystallization and data refinement

Crystallization attempts with the βGly83Cys and βGly83Cys-Phe41Tyr mutants created plate-like crystals that were difficult to loop (see Figure 3.3 and Figure 3.4). After determining stable buffer conditions (see 3.2.2), larger crystals that were slightly thicker than the original crystals were obtained for data collection. Prior to optimization, approximately 1 out of 30 crystals diffracted to at least 3.0 Å, whereas after optimization 1 out of every 16 diffracted to at least 3.0 Å. The hemoglobin unit cell statistics varied significantly from crystal to crystal, regardless of whether or not the crystal came from the same drop or different crystallization conditions. Unit cells contained anywhere from 8 monomers to 64 monomers, depending on the crystal. This created difficulties in obtaining high resolution, complete datasets. Many of the crystals resolution only went between 8-10Å, and only about one out of 16 diffracted better than 4Å.

Crystal hits came from high pH buffers, suggesting that the pH rather than the buffer itself is the main determinant of crystallization. However, resolution and data from the new hits were not improved.

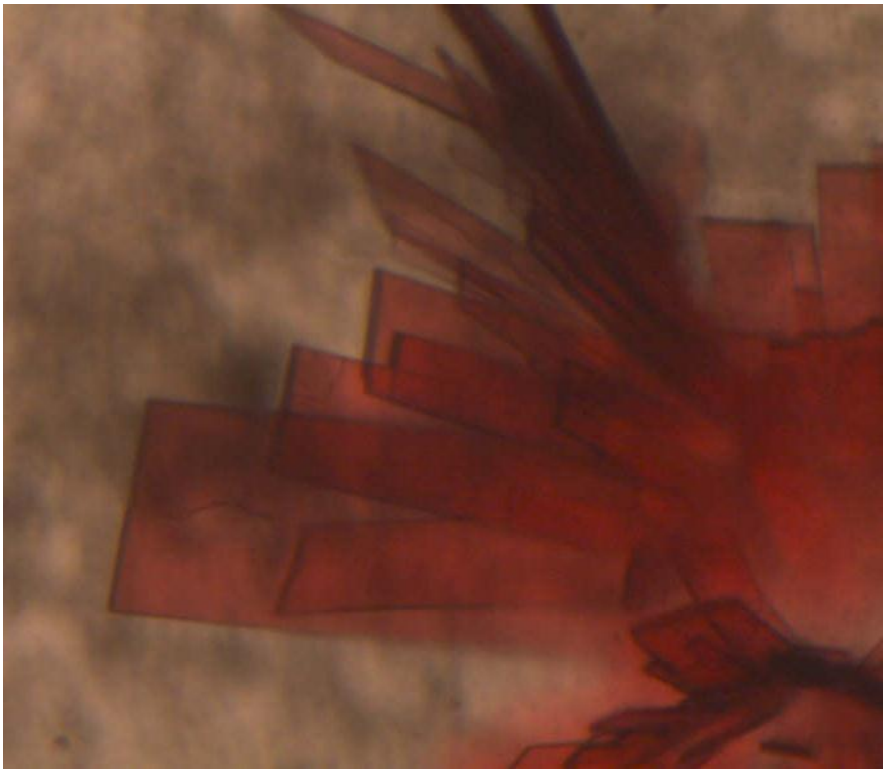

Figure 3.3 βGly83Cys crystals obtained from of 20% w/v PEG 3000, 0.1 M Tris pH 10.0, 0.2 M calcium acetate.

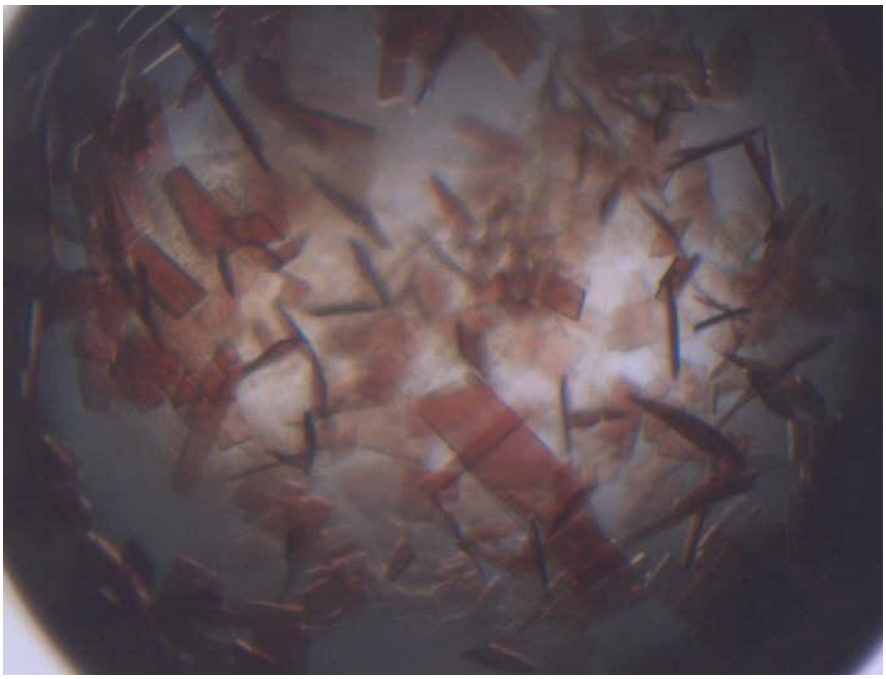

Figure 3.4 βGly83Cys-Phe41Tyr crystals 20% w/v PEG 3000, 0.1 M Tris pH 9.0, 0.2 M calcium acetate.

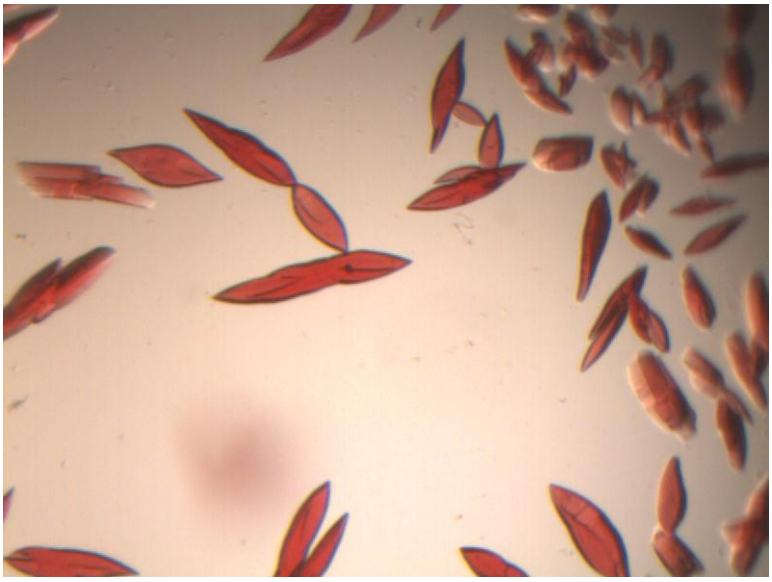

Figure 3.5 αAsn78Cys crystals obtained from 0.17 M lithium sulfate, 85 mM Tris-HCl pH 8.5, 25.5% PEG 4000, and 15% glycerol.

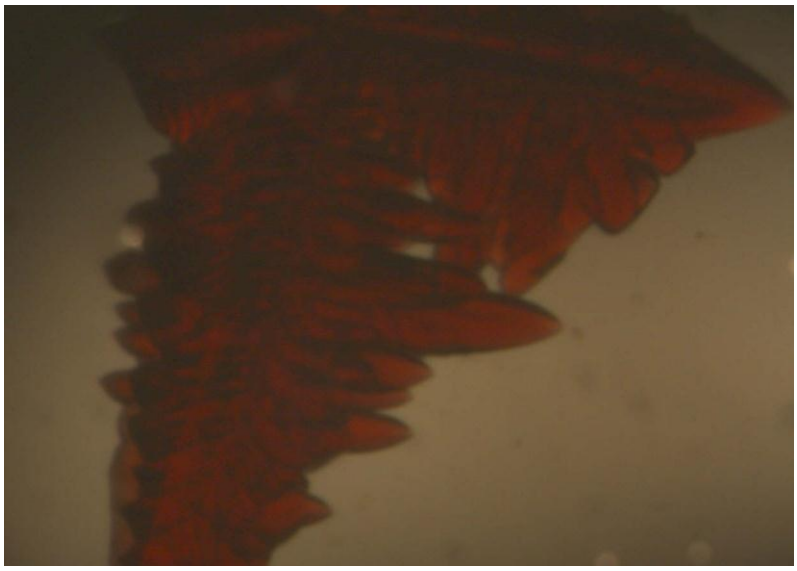

Figure 3.6 βGly83Cys CN in Wizard I 33, consisting of 2.0 M ammonium sulfate, 0.1 M CAPS pH 10.5, 0.2 M lithium sulfate.

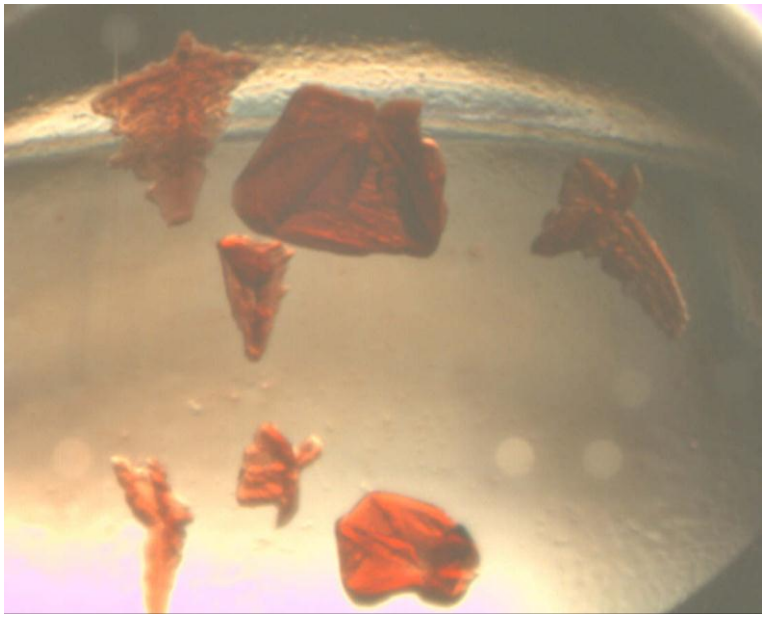

Figure 3.7 βGly83Cys CN in Wiz I 28, consisting of 20% PEG 3000, 0.1 M HEPES pH 7.5, 0.2 M sodium chloride.

3.3.3 Octameric nomenclature

Tetrameric hemoglobin is usually designated with two αβ dimers, $\alpha_1\beta_1$ and $\alpha_2\beta_2$. Octameric hemoglobin necessitates expansion of these designations (see Figure 3.8). For the purposes of discussion, the octamer is said to consist of two tetramers, T1 and T2, combined by a disulfide bridge. Junction between the two tetramers may exist on either the α or the β subunits. Each tetramer contains two αβ dimers. T1 contains $\alpha_1\beta_1$ and $\alpha_2\beta_2$, and T2 contains $\alpha_1'\beta_1'$ and $\alpha_2'\beta_2'$. In the case of the βGly83Cys and βGly83Cys-Phe41Tyr mutants, each asymmetric unit contained two octamers. The second octamer is described as containing tetramers T3 and T4. In cases where differentiating between the two octamers is necessary, T3 consists of $\alpha_3\beta_3$ and $\alpha_4\beta_4$ dimers, and T4 consists of $\alpha_3'\beta_3'$ and $\alpha_4'\beta_4'$ dimers. In places where just the tetramers are compared, for example in comparison with HbA models, the dimer of $\alpha_1'\beta_1'$, $\alpha_3\beta_3$ and $\alpha_3'\beta_3'$ are

considered as the $\alpha_1\beta_1$ dimer. The dimers of $\alpha_2'\beta_2'$, $\alpha_4\beta_4$ and $\alpha_4'\beta_4'$ are considered as the $\alpha_2\beta_2$ dimer.

In both the βGly83Cys and the βGly83Cys-Phe41Tyr, the octamer is produced by disulfide bridges between the β83Cys of β_1 and β83Cys of β_1' and a second disulfide bridge between the β83Cys of β_2 and the β83Cys of β_2'. The orientation of the disulfide bridges in these tetramers creates two tetramers that are related by a pseudo 180° rotation across the disulfide bridge, followed by a secondary rotation of 90° through the $\alpha_1\beta_1$-$\alpha_2\beta_2$ interface (see Figure 3.9).

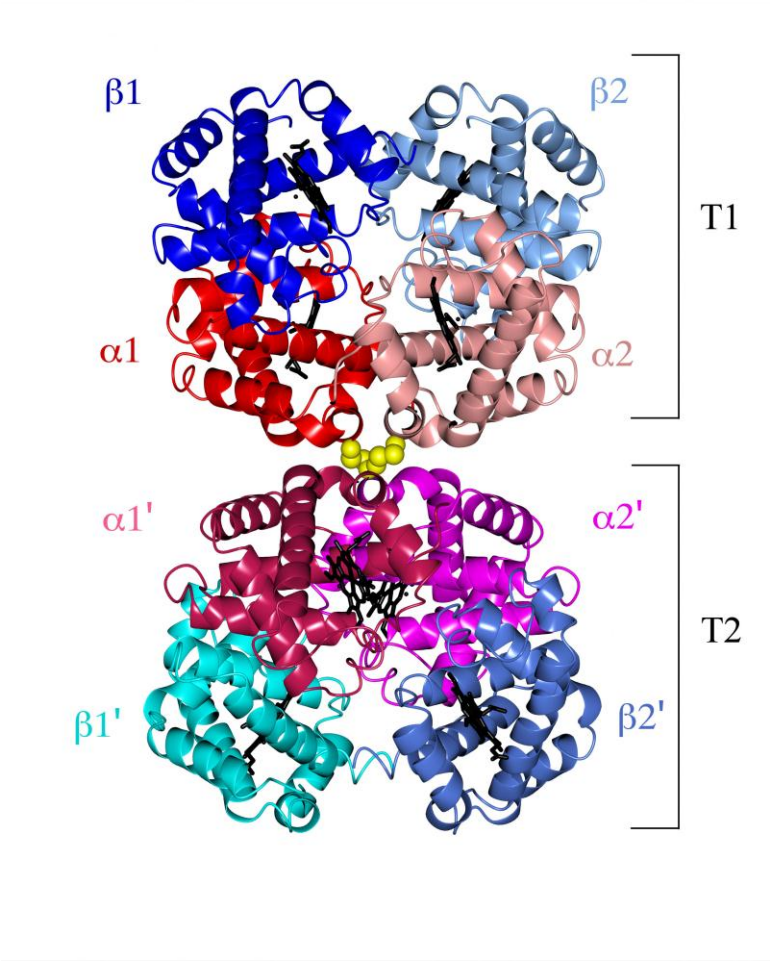

Figure 3.8 Nomenclature of Octameric Hemoglobin

The αAsn78Cys contains two tetramers linked by two separate disulfide bridges, one between α_1 and α_1', the other between α_2 and α_4' (see Figure 3.9 and Figure 3.11). The tetramers of αAsn78Cys are related by a 180° rotation about a non-crystallographic axis going through the two disulfide bridges at the interface of the two tetramers.

The distances between the two tetramers of the αAsn78Cys mutant are much smaller than the βGly83Cys mutants. This provides a structural rationale for the difference in reduction ability between the two mutants (Dr. Chien Ho, personal communication). The βGly83Cys mutants are much more likely to dissociate into tetramers when in the presence of reducing agents than the αAsn78Cys mutant. The structures demonstrate that the increased distance between the Cys Cα atoms allows greater accessibility for the reducing agents (see Table 3.5).

Superposition of the βGly83Cys and αAsn78Cys disulfide bridges demonstrate the differences in disulfide bridge geometry and the relative rotation of the two tetramers. Even when in the same asymmetric unit, the changes at the disulfides alter the relative tetramer conformation (see Figure 3.10). Superimposing one tetramer of the βGly83Cys octamer against the other octamer from the same asymmetric unit (i.e. superimposing tetramer $\alpha_1\beta_1\alpha_2\beta_2$ against $\alpha_3\beta_3\alpha_4\beta_4$) shows that the second tetramer does not superimpose well. The second tetramer may have an RMSD between 4Å and 8Å, depending on the orientation of the first superposition.

Comparison of the mutants show little differences in the tetrameric structures. Superposition of the tetramers to determine the RMSD of the tetramers shows that the tetramers have an RMSD of 0.4 Å (see Table 3.10 and Table 3.11). This RMSD is true regardless of which mutants are compared. RMSD of the βGly83Cys and βGly83Cys-Phe41Tyr octamers also show small RMSD (see Table 3.9). If only the $\alpha_1\beta_1$ dimer is superimposed, the RMSD of the second dimer is around 0.4 Å, showing that the two dimers superimpose very well, regardless of

mutant. Slightly higher RMSD is seen when comparing the βGly83Cys mutants to the αAsn78Cys mutant. The higher RMSD shows that the position of the cysteine bridge slightly changes the tetrameric conformation of the αAsn78Cys mutant.

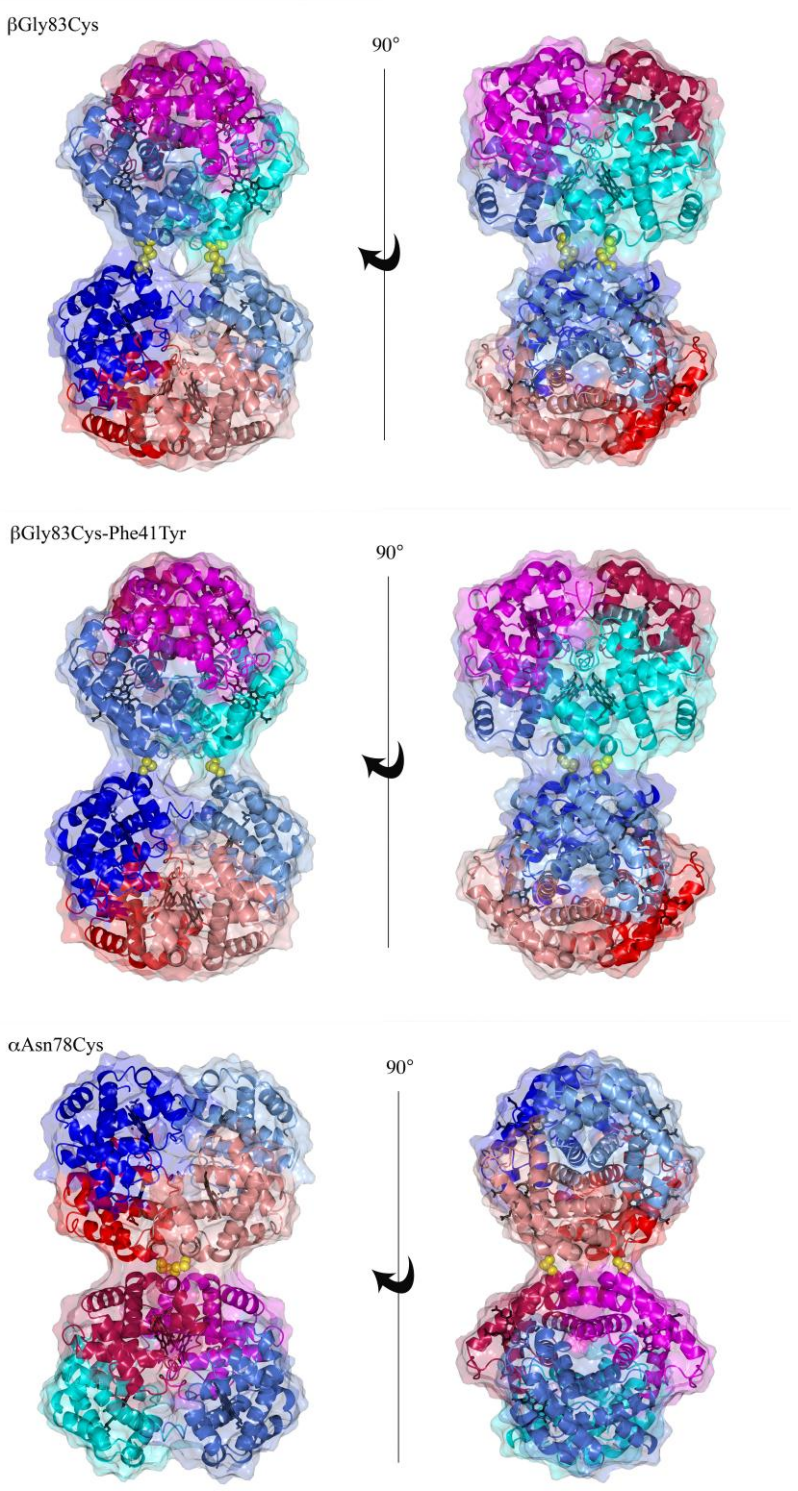

Figure 3.9 Overall octemeric structure of oHb mutants demonstrates the difference in relative tetramer orientation in α and β mutants.α monomers are in shades of red, β monomers are in shades of blue.

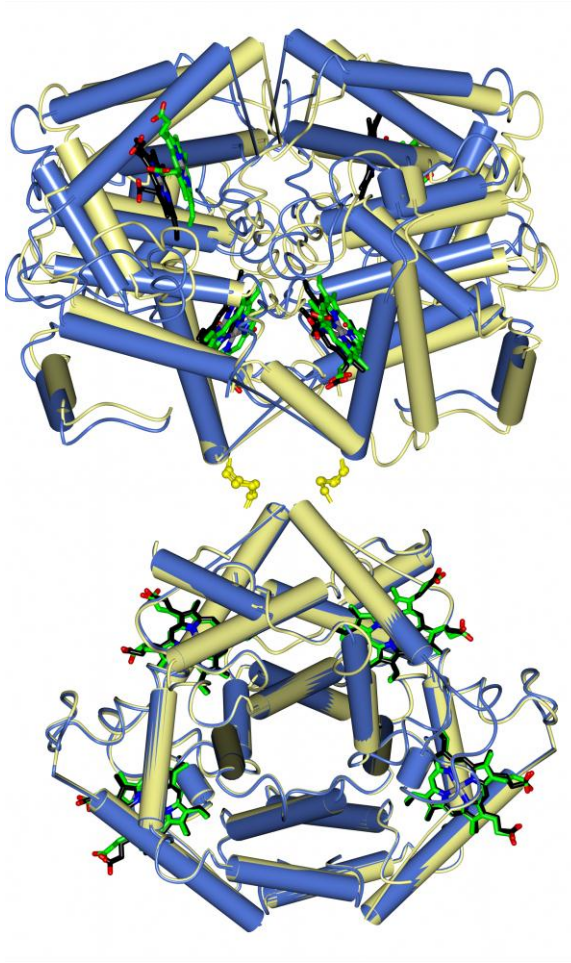

Figure 3.10 Superposition of the two βGly83Cys octamers found in the same asymmetric unit shows changes between the two tetramers are caused by changes in the disulfide bridge geometry.

3.3.4 Stability of the octameric hemoglobin

Octameric hemoglobin was shown to be very stable in the presence of mild reducing agents [34, 119] but the causality was unknown. The octameric structure was presumed to be stabilized by interaction of the two tetramers via two disulfide bridges. The crystal structures of all mutants show distinct electron density between the tetramers indicative of disulfide bridges. The geometry of the disulfide bridges constrains the octamer and the tetramer orientations, with the α

mutant tetramers being at approximately 180° rotation from each other with an axis perpendicular the disulfide bridge. The β mutants have an approximate 90° rotation on an axis parallel to the disulfide bridges, as well as the approximate 180° rotation that allows the mutated cysteines to form covalent bonds.

Comparison of the octameric mutants as dimers and tetramers shows a significant similarity in the relative position of the dimers. Any two mutants have an RMSD of around 0.3 Å for a superimposed α1β1 dimer (see Table 3.6 and Table 3.7). If the RMSD of the α2β2 dimer is determined after α1β1 superposition without superimposing the second dimer, the RMSD is still below 1.0 Å for all mutants (see Table 3.6 and Table 3.8). The two βGly83Cys mutants are the most similar, with RMSD values consistently around 0.3 Å, and both mutants have a higher RMSD when compared to αAsn78Cys.

Table 3.5: Disulfide bridge geometry of the three oHb mutants and rHbCN(βGly83Cys) and rHb (βGly83Cys-Phe41Tyr) demonstrates changes that affect the relative position and reduction of these mutants

	S-S distance (Å)	Cα-Cα distance (Å)	Angle 1 (°)	Angle 2 (°)
rHb (αAsn78Cys)				
α_1-α_1'	2.0	6.1	93	121
α_2-α_2'	2.1	6.0	121	92
rHb (βGly83Cys)				
β_1-β_1'	2.1	5.5	115	104
β_2-β_2'	2.0	5.9	93	106
β_3-β_3'	2.0	5.9	122	106
β_4-β_4'	2.0	5.2	98	112
rHbCN (βGly83Cys)				
β_1-β_1'	2.0	5.5	64	47
β_2-β_2'	2.0	5.3	60	94
rHb (βGly83Cys-Phe41Tyr)				
β_1-β_1'	2.0	5.1	108	99
β_2-β_2'	2.0	5.5	107	110
β_3-β_3'	2.0	5.4	108	110

β_4-β_4'	2.0	5.5	109	108

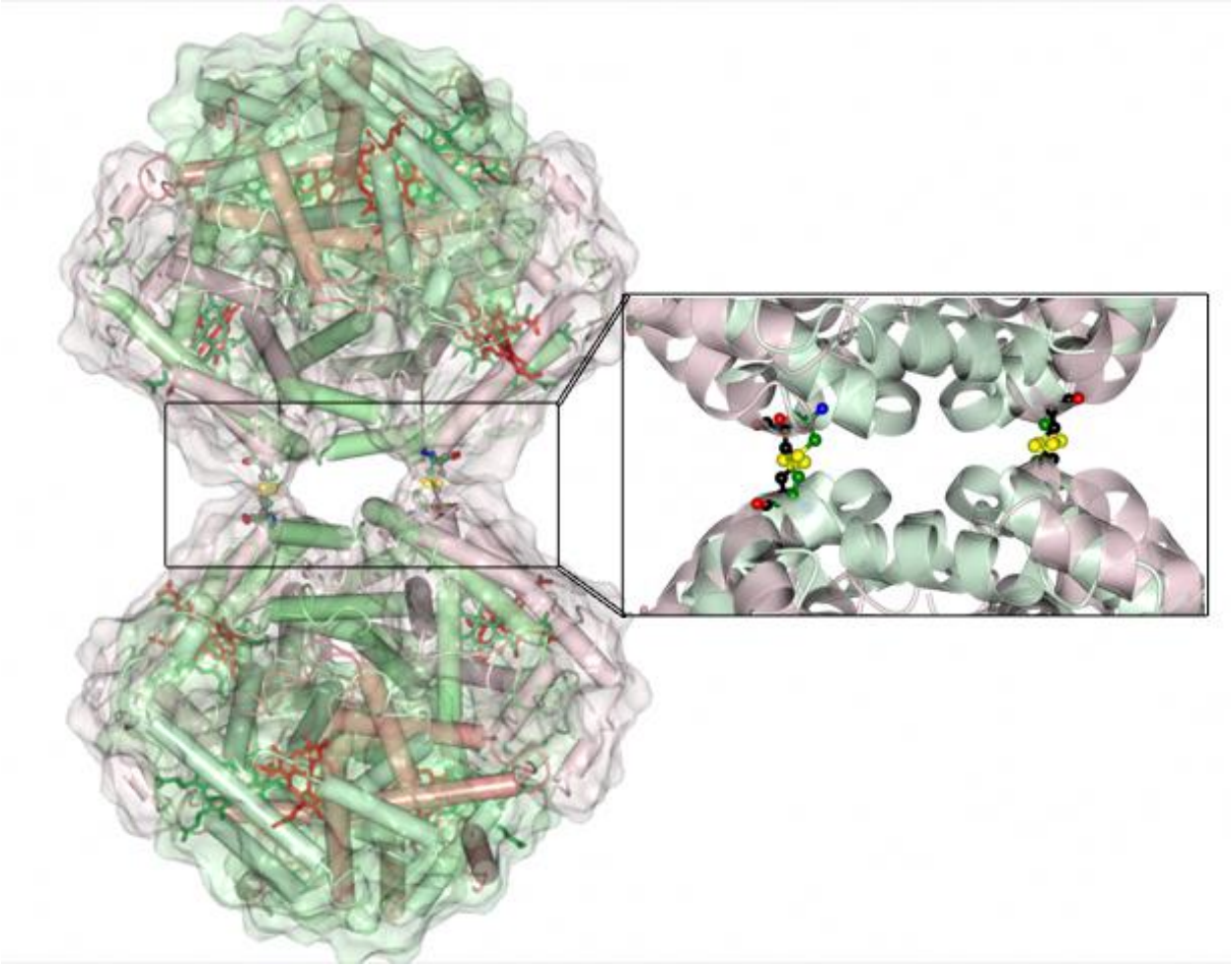

Figure 3.11 Differences in the αAsn78Cys (red) and βGly83Cys (green) disulfide bridges cause changes in the relative orientations of the tetramers. The inset shows an expanded view of the superimposed disulfide bridges.

Table 3.6 RMSD between α1β1 and α2β2. Cα atoms of the $\alpha_1\beta_1$ dimer were superimposed and the RMSD calculated (bottom portion). RMSD of $\alpha_2\beta_2$ Cα atoms were calculated after superposition of the $\alpha_1\beta_1$ dimer, shown in the upper portion of the table. βGly83Cys (G83C), βGly83Cys-Phe41Tyr (F41Y) and αAsn78Cys (N78C) tetramers are shown.

	F41Y T1	F41Y T2	F41Y T3	F41Y T4	G83C T1	G83C T2	G83C T3	G83C T4	N78C T1	N78C T2
F41Y T1	---	0.4	0.3	0.1	0.4	0.4	0.5	0.4	0.8	0.8
F41Y T2	0.1	---	0.2	0.4	0.6	0.4	0.5	0.5	1.1	1.1
F41Y T3	0.1	0.1	---	0.3	0.5	0.4	0.5	0.5	1.0	1.0
F41Y T4	0.0	0.1	0.1	---	0.4	0.4	0.5	0.4	0.9	0.8
G83C T1	0.3	0.3	0.3	0.2	---	0.3	0.3	0.4	0.7	0.6
G83C T2	0.2	0.3	0.3	0.2	0.1	---	0.1	0.3	1.0	0.8
G83C T3	0.3	0.3	0.3	0.3	0.1	0.1	---	0.3	0.9	0.8
G83C T4	0.3	0.3	0.3	0.3	0.1	0.1	0.1	---	0.8	0.8
N78C T1	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2	---	0.7
N78C T2	0.3	0.3	0.3	0.3	0.2	0.3	0.3	0.3	0.3	---

Table 3.7 Average RMSD (Å) of the $\alpha_1\beta_1$ Cα atoms.

	αAsn78Cys	βGly83Cys	βGly83Cys-Phe41Tyr
βGly83Cys-Phe41Tyr			0.1 ±0.02
βGly83Cys		0.1 ±0.02	0.3 ±0.01
αAsn78Cys	0.2	0.3 ±0.02	0.3 ±0.02

Table 3.8 Average RMSD (Å) of the $\alpha_2\beta_2$ Cα atoms after superposition of the $\alpha_1\beta_1$ Cα atoms.

	αAsn78Cys	βGly83Cys	βGly83Cys-Phe41Tyr
βGly83Cys-Phe41Tyr	0.9 ±0.13	0.5 ±0.05	0.3 ±0.11
βGly83Cys	0.8 ±0.10	0.3 ±0.07	
αAsn78Cys	0.5		

Table 3.9 RMSD (Å) of the Cα atoms of the βGly83Cys and βGly83Cys-Phe41Tyr octamers.

	βGly83Cys-Phe41Tyr *O1*	βGly83Cys-Phe41Tyr *O2*	βGly83Cys *O1*	βGly83Cys *O2*
βGly83Cys-Phe41Tyr O1	---			
βGly83Cys-Phe41Tyr O2	0.4	---		
βGly83Cys O1	0.8	1.0	---	
βGly83Cys O2	1.0	0.8	1.2	---

Table 3.10 RMSD (Å) of the tetramers of each of the oxidized mutants after superposition of the tetramer Cα atoms. βGly83Cys (G83C), βGly83Cys-Phe41Tyr (F41Y) and αAsn78Cys (N78C).

	F41Y T1	F41Y T2	F41Y T3	F41Y T4	G83C T1	G83C T2	G83C T3	G83C T4	N78C T1	N78C T2
F41Y T1	---									
F41Y T2	0.2	---								
F41Y T3	0.1	0.1	---							
F41Y T4	0.0	0.2	0.1	---						
G83C T1	0.3	0.3	0.3	0.3	---					
G83C T2	0.3	0.3	0.3	0.4	0.2	---				
G83C T3	0.3	0.3	0.3	0.4	0.2	0.1	---			
G83C T4	0.3	0.3	0.3	0.3	0.2	0.2	0.2	---		
N78C T1	0.4	0.5	0.5	0.4	0.4	0.4	0.4	0.4	---	
N78C T2	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.4	0.5	---

Table 3.11 Average RMSD (Å) of the Cα atoms after superposition of the mutant tetramers.

	βGly83Cys-Phe41Tyr	βGly83Cys	αAsn78Cys
βGly83Cys-Phe41Tyr	0.1 ±0.05		
βGly83Cys	0.3 ±0.02	0.2 ±0.04	
αAsn78Cys	0.4 ±0.04	0.4 ±0.04	0.3

3.3.5 Comparison of octameric mutants and previously-determined models

A large question about the relative reactivity between the octameric mutants and previously-identified hemoglobin models has to do with the large overall structural changes that occur with the addition of the mutations. As of May 2012, over 240 models of human hemoglobin were present in the PDB, many of which are mutants. Table 3.12 shows models that were used to compare the octameric oHb. These models are more recent than earlier models that showed oxidized Hb, and are more likely to be refined with programs that have more physiologically relevant restraints. Using these models should prevent problems inherent in comparing apples to oranges, giving inconsistent comparison lengths.

Table 3.12 PDB ID and Characteristics of HbA Models Used for Comparison

PDB ID	Characteristics
1A3N [12]	1.25 Å T state
1GZX [10]	2.10 Å Oxygen-bound T state
2DN3 [2]	1.25 Å Carbon monoxy-bound R state
2DN1 [2]	1.25 Å Oxygen-bound R state
1BBB [17]	1.7 Å Carbon monoxy-bound R2 state, low salt
1MKO [6]	2.18 Å Carbon monoxy-bound RR2 state
1YZI [6]	2.07 Å Carbon monoxy-bound R3 state

There are multiple ways to compare the various models. One way is using root mean square deviation (RMSD), which takes the distances between comparable atoms in the two models and averages the distances. A larger RMSD means that the sections have a larger conformational difference. RMSD can be affected by outlying sections of a model. Previous research has shown that Hb changes that are a consequence of ligand binding can be examined by RMSD and described by a relative rotation of the $\alpha_2\beta_2$ dimer relative to the $\alpha_1\beta_1$ dimer. A comparison of the required rotations and RMSD shows that all of the models are closest to the R2-state model PDB ID 1BBB. Table 3.13 to Table 3.18 show the RMSD and angles determined by comparing the octameric structures against previously-published HbA structures. All models are close to the R2 state shown in PDB ID 1BBB regardless of crystallization condition. While the PEG-crystallized aquamet mutants would be expected to be in the R2 state, this is the first high-salt structure that is close to the R2 state. The results of the high salt structure suggest that this conformation is not entirely determined by the crystallization conditions, and that the disulfide bridge is constraining the tetramers into the R2 conformation. Table 3.19 to Table 3.24 give the distances between residues of the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers. The interactions that are listed are support the R-state interpretation and demonstrate that the disulfide bridges do not alter the distances between the $\alpha\beta$ dimers. The movement of the tetramers is due to the disulfide bridges, which appear to alter their geometry to accommodate any strain on the tetramers.

Table 3.13 Comparison of cyanide-bound high salt βGly83Cys with published HbA structures. Dimers were superimposed separately to obtain the RMSD (Å). In places where only one dimer is listed, the second dimer is obtainable by symmetry relations and would give the same RMSD as the $\alpha_1\beta_1$ dimer.

	2DN3 [159]	1BBB [160] $\alpha_1\beta_1$	1BBB $\alpha_2\beta_2$	1GZX [161] $\alpha_1\beta_1$	1GZX $\alpha_2\beta_2$	1YZI [6]	1A3N [12] $\alpha_1\beta_1$	1A3N $\alpha_2\beta_2$	1MKO [6] $\alpha_1\beta_1$	1MKO $\alpha_2\beta_2$	2DN1 [159]
$\alpha_1\beta_1$	0.93	0.34	0.31	1.19	1.17	1.16	1.23	1.25	0.71	0.53	0.97
$\alpha_2\beta_2$	0.89	0.40	0.32	1.13	1.11	1.15	1.17	1.19	0.70	0.50	0.92
$\alpha_1'\beta_1'$	0.91	0.38	0.34	1.16	1.14	1.14	1.20	1.22	0.71	0.54	0.93
$\alpha_2'\beta_2'$	0.92	0.37	0.30	1.18	1.17	1.17	1.22	1.23	0.73	0.52	0.96
	0.91 ±0.02	0.37 ±0.03	0.32 ±0.02	1.16 ±0.01	1.15 ±0.03	1.16 ±0.01	1.20 ±0.03	1.22 ±0.03	0.72 ±0.01	0.52 ±0.02	0.95 ±0.02
		0.34±0.03		1.16±0.02			1.21±0.03		0.62±0.11		

Table 3.14 βGly83Cys comparison against various tetrameric HbA models. The entire tetramer was superimposed to obtain the tetramer RMSD. The $\alpha_1\beta_1$ dimer was superimposed to obtain the $\alpha_1\beta_1$ RMSD, after which the RMSD of the nonsuperimposed $\alpha_2\beta_2$ dimer and the angle necessary to superimpose the second dimer were determined.

Tetramer	Tetramer RMSD (Å)	$\alpha_1\beta_1$ RMSD (Å)	$\alpha_2\beta_2$ RMSD (Å)	$\alpha_2\beta_2$ angle (°)	$\alpha_2\beta_2$ RMSD (Å)
1A3N					
T1	3.4	1.2	8.5	25.6	1.3
T2	3.5	1.2	8.8	26.2	1.3
T3	3.4	1.2	8.7	25.7	1.3
T4	3.4	1.2	8.6	25.5	1.2
1GZX					
T1	3.6	1.2	9.1	27.0	1.2
T2	3.7	1.2	9.4	27.6	1.2
T3	3.6	1.2	9.3	27.1	1.2
T4	3.6	1.2	9.3	26.9	1.2
2DN3					
T1	1.6	1.0	3.8	8.2	1.0
T2	1.6	1.0	4.0	8.5	1.0
T3	1.6	1.0	4.0	8.0	1.0
T4	1.6	1.0	3.9	8.1	1.0
2DN1					
T1	1.7	1.0	4.1	8.9	1.0
T2	1.7	1.0	4.3	9.1	1.0
T3	1.4	1.0	4.3	8.6	1.0
T4	1.7	1.0	4.2	8.7	1.0
1BBB					
T1	0.6	0.5	0.9	6.0	0.4
T2	0.5	0.4	0.7	5.4	0.4
T3	0.5	0.4	0.8	5.8	0.4
T4	0.6	0.4	0.9	6.1	0.4
1MKO					
T1	1.0	0.8	2.0	5.3	0.6
T2	1.0	0.8	2.2	5.4	0.5
T3	0.9	0.8	2.2	5.0	0.6
T4	0.9	0.8	2.1	5.2	0.6
1YZI					
T1	2.5	1.2	5.6	17.1	1.1
T2	2.5	1.1	5.7	17.4	1.1
T3	2.5	1.2	5.7	16.9	1.1

T4	2.4	1.1	5.6	16.9	1.2

Table 3.15 βGly83Cys-Phe41Tyr comparison against various HbA models. The entire tetramer was superimposed to obtain the tetramer RMSD. The $\alpha_1\beta_1$ dimer was superimposed to obtain the $\alpha_1\beta_1$ RMSD, after which the RMSD of the nonsuperimposed $\alpha_2\beta_2$ dimer and the angle necessary to superimpose the second dimer were determined.

Tetramer	Tetramer RMSD (Å)	$\alpha_1\beta_1$ RMSD (Å)	$\alpha_2\beta_2$ RMSD (Å)	$\alpha_2\beta_2$ angle (°)	$\alpha_2\beta_2$ RMSD (Å)
1A3N					
T1	3.4	1.2	8.6	27.8	1.3
T2	3.5	1.2	8.9	28.4	1.2
T3	3.5	1.2	8.8	28.3	1.2
T4	3.5	1.2	8.7	28.0	1.3
1GZX					
T1	3.6	1.2	9.3	29.2	1.2
T2	3.7	1.2	9.6	29.8	1.1
T3	3.6	1.2	9.5	29.7	1.1
T4	3.6	1.2	9.3	29.3	1.1
2DN3					
T1	1.6	1.0	3.9	8.7	1.0
T2	1.7	1.0	4.2	9.3	1.0
T3	1.7	1.0	4.1	9.3	1.0
T4	1.6	1.0	4.0	8.8	1.0
2DN1					
T1	1.7	1.0	4.2	9.3	1.0
T2	1.8	1.0	4.5	10.0	1.0
T3	1.8	1.0	4.4	9.9	1.0
T4	1.7	1.0	4.3	9.4	1.0
1BBB					
T1	0.5	0.4	0.9	3.8	0.4
T2	0.5	0.4	0.9	3.2	0.4
T3	0.5	0.4	1.0	3.6	0.4
T4	0.5	0.4	0.9	3.6	0.4
1MKO					
T1	0.9	0.8	2.0	5.5	0.6
T2	1.0	0.8	2.4	6.1	0.6
T3	1.0	0.8	2.3	6.1	0.5
T4	1.0	0.8	2.1	5.6	0.6
1YZI					
T1	2.5	1.2	5.6	17.8	1.2
T2	2.6	1.2	5.9	18.5	1.2

T3	2.6	1.2	5.8	18.4	1.2
T4	2.5	1.2	5.6	18.0	1.2

Table 3.16 αAsn78Cys comparison with various HbA models. The entire tetramer was superimposed to obtain the tetramer RMSD. The $\alpha_1\beta_1$ dimer was superimposed to obtain the $\alpha_1\beta_1$ RMSD, after which the RMSD of the nonsuperimposed $\alpha_2\beta_2$ dimer and the angle necessary to superimpose the second dimer were determined.

Tetramer	Tetramer RMSD (Å)	$\alpha_1\beta_1$ RMSD (Å)	$\alpha_2\beta_2$ RMSD (Å)	$\alpha_2\beta_2$ angle (°)	$\alpha_2\beta_2$ RMSD (Å)
1A3N					
T1	3.2	1.2	7.9	24.8	1.2
T2	3.4	1.2	8.2	26.0	1.2
1GZX					
T1	3.3	1.1	8.5	26.2	1.1
T2	3.5	1.1	8.8	27.3	1.1
2DN3					
T1	1.4	0.9	3.2	8.1	1.0
T2	1.5	1.0	3.3	7.4	0.9
2DN1					
T1	1.5	1.0	3.5	8.3	1.0
T2	1.6	1.0	3.6	7.8	1.0
1BBB					
T1	0.6	0.4	1.2	6.8	0.4
T2	0.6	0.5	1.2	5.6	0.4
1MKO					
T1	0.8	0.7	1.5	5.3	0.5
T2	0.9	0.8	1.5	4.3	0.5
1YZI					
T1	2.5	1.2	5.3	16.8	1.2
T2	2.4	1.2	5.1	16.4	1.2

Table 3.17 Average Cα RMSD (Å) of the tetramers

	1A3N	1GZX	2DN1	2DN3	1MKO	1BBB	1YZI
βGly83Cys	3.4±0.0	3.6±0.0	1.6±.1	1.6±0.0	0.9±0.0	0.5±0.0	2.5±0.0
βGly83Cys-Phe41Tyr	3.5±0.0	3.6±0.0	1.8±0.1	1.7±.1	1.0±0.0	0.5±0.0	2.6±0.0
αAsn78Cys	3.5±0.7	3.4±2	1.5±0.1	1.4±.1	0.8±0.0	0.6±0.0	2.4±0.1

Table 3.18 Average angle necessary to superimpose $\alpha_2\beta_2$ dimers after superposition of $\alpha_1\beta_1$ dimers

	1A3N	1GZX	2DN1	2DN3	1MKO	1BBB	1YZI
βGly83Cys	25.8±0.3	27.1±0.3	8.8±0.2	8.2±0.2	5.2±0.2	5.8±0.3	17.1±0.2
βGly83Cys-Phe41Tyr	28.1±0.3	29.5±.3	9.6±0.4	9.0±0.4	5.8±0.4	3.6±0.2	18.2±0.4
αAsn78Cys	25.4±1.2	26.8±1.2	8.1±0.5	7.8±0.7	4.8±1.0	6.2±1.3	16.6±0.4

Table 3.19 Hydrogen bonding contacts (Å) between $\alpha_1\beta_1$ dimers for αAsn78Cys dimers

α atom	β atom	α_1-β_1	α_2-β_2	α_1'-β_1'	α_2'-β_2'
Arg31 NH	Gln127 OE1	2.7	2.9	3.0	2.9
Arg31 NH	Phe122 O	3.2	3.2	3.1	3.1
Arg31 NH	The123 O	3.8	3.8	3.9	3.7
Arg31 NH	Pro124 N	3.6	3.5	3.6	3.5
Arg31 NH	Phe122 O	2.9	2.9	2.9	2.8
Ser35 OG	Gln127 NE2	3.9	3.8	3.7	3.8
Ser35 OG	Ala128 N	3.6	3.5	3.5	3.4
His103 NE2	Asn108 OD1	3.6	3.7	3.7	3.6
His103 NE2	Gln131 OE1	2.8	2.8	2.8	2.7
His103 NE2	Gln127 NE2	4.0	3.9	3.9	4.0
Cys104 SG	Gln127 NE2	3.9	3.8	3.8	3.8
Ala110 O	Ala115 O	3.9	3.9	4.0	3.9
Pro114 O	His116 NE2	2.8	2.7	2.9	2.8
Phe117 O	Arg30 NH1	3.4	3.4	3.4	3.4
Phe117 O	Arg30 NH2	3.0	2.8	2.8	3.0
Phe117 O	His116 NE2	3.4	3.3	3.5	3.6
Thr118 O	Arg30 NH	3.5	3.4	3.5	3.5
Pro119 N	Arg30 NH1	3.7	3.6	3.7	3.7
Pro119 N	Arg30 NH2	3.7	3.6	3.7	3.7
His112 ND1	Arg30 NH	3.0	2.9	2.9	3.1
Asp126 OD	Tyr35 OH	3.4	3.4	3.4	3.4

Table 3.20 Hydrogen bonding contacts (Å) between α_1-β_2 dimers in αAsn78Cys mutant

α atom	β atom	α_1-β_2	α_2-β_1	α_1'-β_2'	α_2'-β_1'
Thr41 O	Arg40 NH	3.3	3.1	3.0	3.4
Arg92 O	Trp37 O	3.8	3.7	3.8	3.9
Asp94 OD1	Trp37 NE1	3.9	3.9	3.7	3.9
Asp94 OD2	Trp37 NE1	3.4	3.5	3.3	3.5
Asp94 OD2	Asn102 OD1	3.6	3.6	3.8	3.7
Asp94 OD2	Asn102 ND2	2.8	2.7	3.0	2.9

Table 3.21 Hydrogen bonding contacts (Å) between the α_1 and β_1 subunits in the βGly83Cys mutant

α atom	β atom	α_1-β_1	α_2-β_2	α_1'-β_1'	α_2'-β_2'	α3-β3	α4-β4	α3′-β3′	α4′-β4′
Arg31 NE	Gln127 OE1	4.0	3.7	3.8	3.8	3.7	3.9	3.8	3.9
Arg31 NH	Gln127 OE1	2.8	2.6	2.7	2.7	2.7	2.8	2.7	2.7
Arg31 NH	Phe122 O	3.3	3.4	3.3	3.3	3.3	3.3	3.3	3.4
Arg31 NH	The123 O	4.0	3.8	3.8	3.9	3.7	--	3.8	--
Arg31 NH	Pro124 N	3.5	3.5	3.4	3.5	3.4	3.5	4.0	3.5
Arg31 NH	Phe122 O	2.8	2.8	2.7	2.8	2.7	2.8	2.8	3.0
Ser35 OG	Gln127 NE2	3.8	3.9	4.0	3.8	3.9	3.8	3.9	3.9
Ser35 OG	Ala128 N	3.6	3.5	3.4	3.5	3.5	3.5	3.5	3.7
Lys99 NZ	Glu101 OE	3.8	3.5	3.6	3.6	3.8	3.5	3.7	3.8
His103 NE2	Asn108 OD1	3.9	3.8	3.8	3.8	3.8	3.9	3.8	3.9
His103 NE2	Gln131 OE1	2.8	2.8	2.7	2.8	2.8	2.7	2.8	2.8
Cys104 SG	Gln127 NE2	3.8	3.7	3.6	3.7	3.7	3.7	3.8	3.9
Ala110 O	Ala115 O	3.7	3.7	3.7	3.8	3.6	3.8	3.7	3.6
Pro114 O	His116 NE2	2.8	2.8	2.8	2.9	2.8	2.9	2.9	2.8
Phe117 O	Arg30 NH1	2.8	2.9	2.7	2.9	2.9	2.9	2.9	2.8
Phe117 O	Arg30 NH2	3.4	3.3	3.4	3.4	3.3	3.5	3.5	3.4
Phe117 O	His116 NE2	3.6	3.7	3.7	3.7	3.6	3.8	3.6	3.6
Thr118 O	Arg30 NH	3.6	3.6	3.5	3.6	3.5	3.8	3.7	3.5
Pro119 N	Arg30 NH1	3.6	3.7	3.7	3.7	3.7	3.8	3.7	3.6
Pro119 N	Arg30 NH2	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.4
His112 ND1	Arg30 NH	3.1	3.2	3.2	3.1	3.2	3.8	3.1	3.2
Asp126 OD	Tyr35 OH	3.3	3.1	3.1	3.3	3.3	3.2	3.2	3.2

Table 3.22 Hydrogen bonding (Å) distances between the α_1 and β_2 monomers in βGly83Cys mutant

α atom	β atom	α_1-β_1	α_2-β_2	α_1'-β_1'	α_2'-β_2'	α3-β3	α4-β4	α3′-β3′	α4′-β4′
Thr41 OG1	Arg40 NH	3.8	4.0	3.8	3.8	3.9	3.6	3.9	3.1
Thr41 O	Arg40 NH	3.6	3.4	3.5	3.5	3.5	3.4	3.5	3.0
Tyr42 OH	Arg40 NH	3.7	3.9	3.7	3.9	3.7	4.0	3.9	4.0
Arg92 NH1	Gln39 OE1	3.6	3.5	3.6	3.5	3.6	3.3	3.5	3.0
Arg92 NH	Glu43 OE	3.9	3.7	3.8	3.8	3.9	--	3.8	--
Arg92 O	Trp37 O	3.7	3.6	3.7	3.8	3.9	3.8	3.7	3.7
Asp94 OD1	Trp37 NE1	3.9	3.9	3.9	3.9	3.9	3.8	3.8	3.6
Asp94 OD2	Trp37 NE1	3.7	3.6	3.8	3.6	3.7	3.5	3.6	3.3
Asp94 OD1	Asn102 OD1	3.9	3.6	3.8	3.8	3.8	3.8	3.7	3.7
Asp94 OD2	Asn102 ND2	2.9	2.7	2.8	3.0	2.8	3.1	2.8	3.0

Table 3.23 Hydrogen bonding distances (Å) between α_1 and β_1 monomers in the βGly83Cys-Phe41Tyr mutant

α atom	β atom	α_1-β_1	α_2-β_2	α_1'-β_1'	α_2'-β_2'	α_3-β_3	α_4-β_4	α_3'-β_3'	α_4'-β_4'
Arg31 NE	Gln127 OE1	3.9	4.0	3.9	4.0	3.9	3.9	3.9	3.9
Arg31 NH	Gln127 OE1	2.8	2.9	2.9	3.0	2.9	2.9	2.9	2.9
Arg31 NH	Phe122 O	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
Arg31 NH	The123 O	3.6	3.8	3.6	3.6	3.7	3.6	3.7	3.6
Arg31 NH	Pro124 N	3.6	3.6	3.6	3.5	3.6	3.4	3.6	3.5
Arg31 NH	Phe122 O	2.8	2.8	2.8	2.8	2.8	2.7	2.8	2.8
Ser35 OG	Gln127 NE2	3.9	3.8	3.8	3.8	3.7	3.9	3.7	3.8
Ser35 OG	Ala128 N	3.4	3.5	3.4	3.5	3.5	3.4	3.5	3.5
Lys99 NZ	Glu101 OE								
His103 NE2	Asn108 OD1	3.8	3.9	3.9	4.0	3.9	3.9	3.9	3.8
His103 NE2	Gln131 OE1	2.8	2.8	2.8	2.8	2.9	2.7	2.7	2.7
Cys104 SG	Gln127 NE2	3.8	3.9	3.8	3.9	3.8	3.8	3.7	3.8
Ala110 O	Ala115 O	4.0	--	4.0	3.9	4.0	3.9	--	--
Pro114 O	His116 NE2	2.7	2.8	2.7	2.7	2.8	2.7	2.8	2.7
Phe117 O	Arg30 NH1	3.5	3.4	3.5	3.5	3.5	3.5	3.5	3.4
Phe117 O	Arg30 NH2	2.8	2.7	2.7	2.7	2.7	2.7	2.8	2.7
Phe117 O	His116 NE2	3.8	3.8	3.6	3.5	3.7	3.5	3.8	3.8
Thr118 O	Arg30 NH	3.6	3.5	3.4	3.5	3.6	3.4	3.6	3.6
Pro119 N	Arg30 NH1	3.5	3.6	3.5	3.5	3.6	3.4	3.6	3.6
Pro119 N	Arg30 NH2	3.8	3.9	3.7	3.7	3.8	3.6	3.9	3.9
His112 ND1	Arg30 NH	3.2	3.2	3.0	3.0	3.0	3.1	3.1	3.1
Asp126 OD	Tyr35 OH	3.1	3.2	3.1	3.2	3.2	3.1	3.2	3.2

Table 3.24 Hydrogen bonding distances (Å) between α_1 and β_2 monomers in the βGly83Cys-Phe41Tyr mutant

α atom	β atom	α_1-β_1	α_2-β_2	α_1'-β_1'	α_2'-β_2'	α_3-β_3	α_4-β_4	α_3'-β_3'	α_4'-β_4'
Thr41 OG1	Arg40 NH	3.5	3.9	3.4	3.2	3.6	3.5	3.3	3.4
Tyr42 OH	Arg40 NE	3.8	3.7	3.6	3.6	3.7	3.5	3.8	3.5
Arg92 NH1	Gln39 OE1	3.1	3.0	3.0	3.0	2.9	3.0	3.0	3.1
Arg92 NH	Glu43 OE	3.5	3.6	3.7	3.7	3.7	3.7	3.5	3.6
Arg92 O	Trp37 O	3.7	3.6	3.6	3.6	3.6	3.6	3.8	3.7
Asp94 OD1	Trp37 NE1	4.0	4.0	3.9	3.9	3.9	3.9	--	3.9
Asp94 OD2	Trp37 NE1	3.8	3.7	3.8	3.8	3.6	3.8	3.8	3.7
Asp94 OD1	Asn102 OD1	4.0	3.9	--	--	3.9	4.0	4.0	3.9
Asp94 OD2	Asn102 ND2	2.9	2.8	2.9	2.9	2.8	2.8	2.9	2.8
Asp94 OD	Tyr41 OH	3.0	2.9	2.9	2.8	3.0	2.8	3.1	3.0

Neither the βGly83Cys-Phe41Tyr nor the βGly83Cys structures contain the interactions between βHis146 and βAsp94 that are indicative of the low-affinity T state. While the lack of salt bridge is expected in the liganded βGly83Cys-Phe41Tyr model, the lack of interaction between these two residues in the βGly83Cys model suggests that the disulfide bridge between the two tetramers breaks the interaction, bringing the structure closer to the R state. The positioning of the terminal regions stabilizes a liganded octameric hemoglobin rather than an unliganded, increasing the affinity for CO or O_2. Restraints put on the protein by lattice contacts may also contribute some of the stabilization force. The lack of interaction explains the decrease in cooperativity of oxygen binding. Mutants where the terminal histidine has been deleted show less cooperativity than HbA, presumably because the interactions between βHis146 and βAsp94 stabilize the T state. While the βHis146-βAsp94 interaction is not seen in the αAsn78Cys mutant, a compensatory interaction is seen between the oxygen of βAsp94 and the βLys144

nitrogen (2.9 Å). The hydrogen bond is not seen in the βGly83Cys structure, and instead a weak interaction (4.2 Å) between the oxygen of βGlu90 and nitrogen of βLys144 is present.

The βGly83Cys-Phe41Tyr mutant contains an additional hydrogen bonding interaction between β_1Tyr41 and α_2Asp94. The hydroxyl group also interacts with the α main chain, hydrogen bonding with the βAsp99 nitrogen and the βHis97 carbonyl. The above listed interactions would affect the position of βPhe42 at the heme pocket, thus affecting ligand binding.

3.3.6 Comparison of Heme Pockets

The heme pockets of the α chain show more consistency in the relative placement of the histidines. See Table 3.25 and Table 3.26 for the geometry of the ligand and histidines that coordinate the heme iron. The αAsn78Cys model has the distal histidine closer to the center of the monomer. The βGly83Cys-Phe41Tyr distal histidines are pushed the furthest from the center of the heme group, while the βGly83Cys histidines adopt a conformation in between that of the αAsn78Cys and βGly83Cys-Phe41Tyr histidine positions. The differences in the distal histidine positions in the βGly83Cys-Phe41Tyr may be due to the presence of hydroxyl ligand. The ligand coordinates between the iron and the nitrogen of the distal histidine and increases the distance between the two when present. The αAsn78Cys structure has the same change in histidine position depending on heme ligand state, but shows a conformation closer to the condensed R state seen in previous crystal structures.

The electron density of the oxidized octameric mutants shows that all of the mutants have only some heme groups bound to ligands. Figure 3.12 to Figure 3.14 show representative images

of the electron density of the α and β heme pocket. The images show the 2Fo-Fc electron density at a contour level of 3σ. The α heme pockets contain hydroxyl ligand the majority of the time, with only the βGly83Cys mutant containing α monomers without ligand present. The β heme pockets were much more variable, containing a mixture of liganded and unliganded pockets.

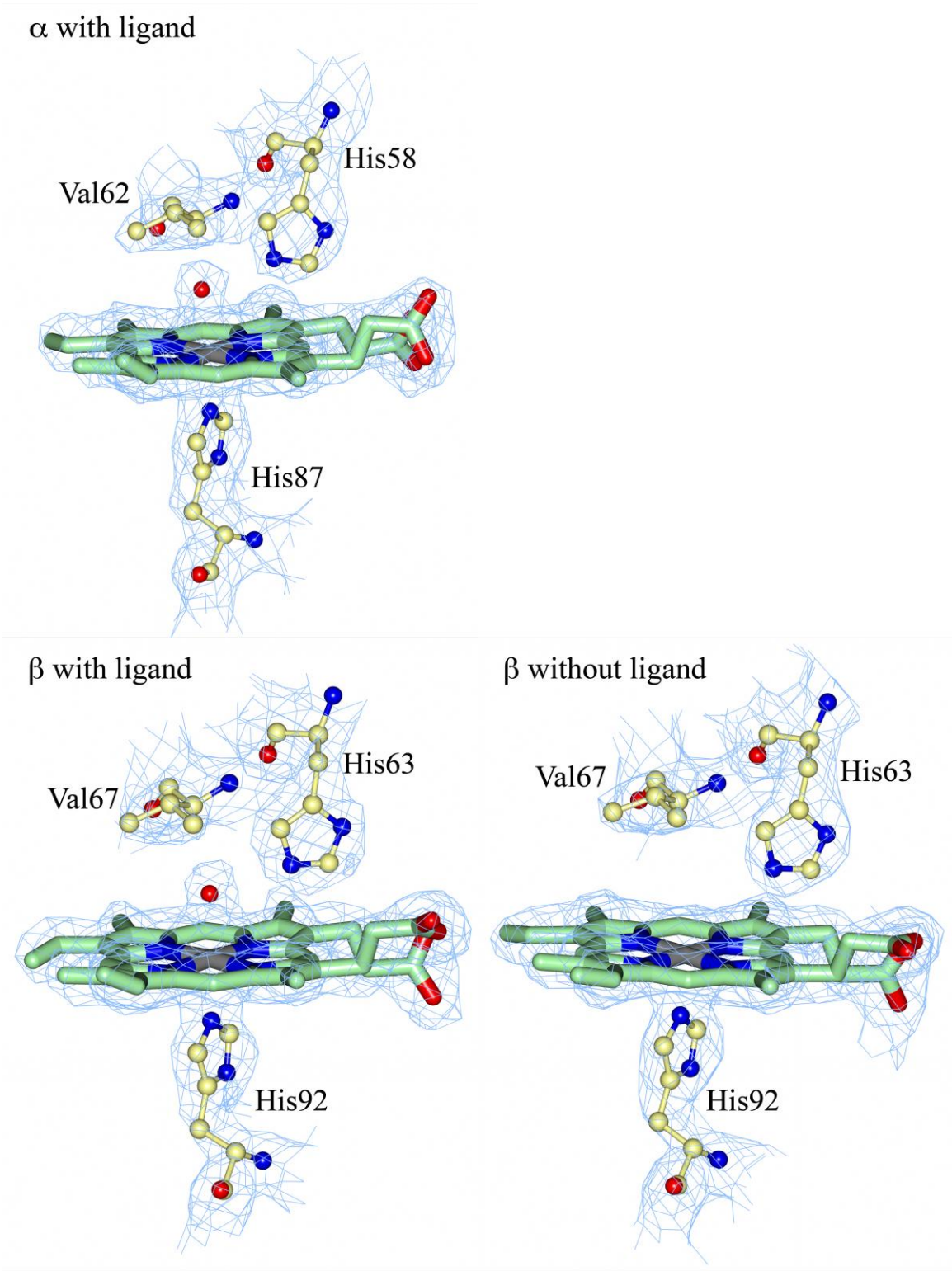

Figure 3.12 Electron density of representative heme pockets demonstrating the aquamet form of βGly83Cys-Phe41Tyr.

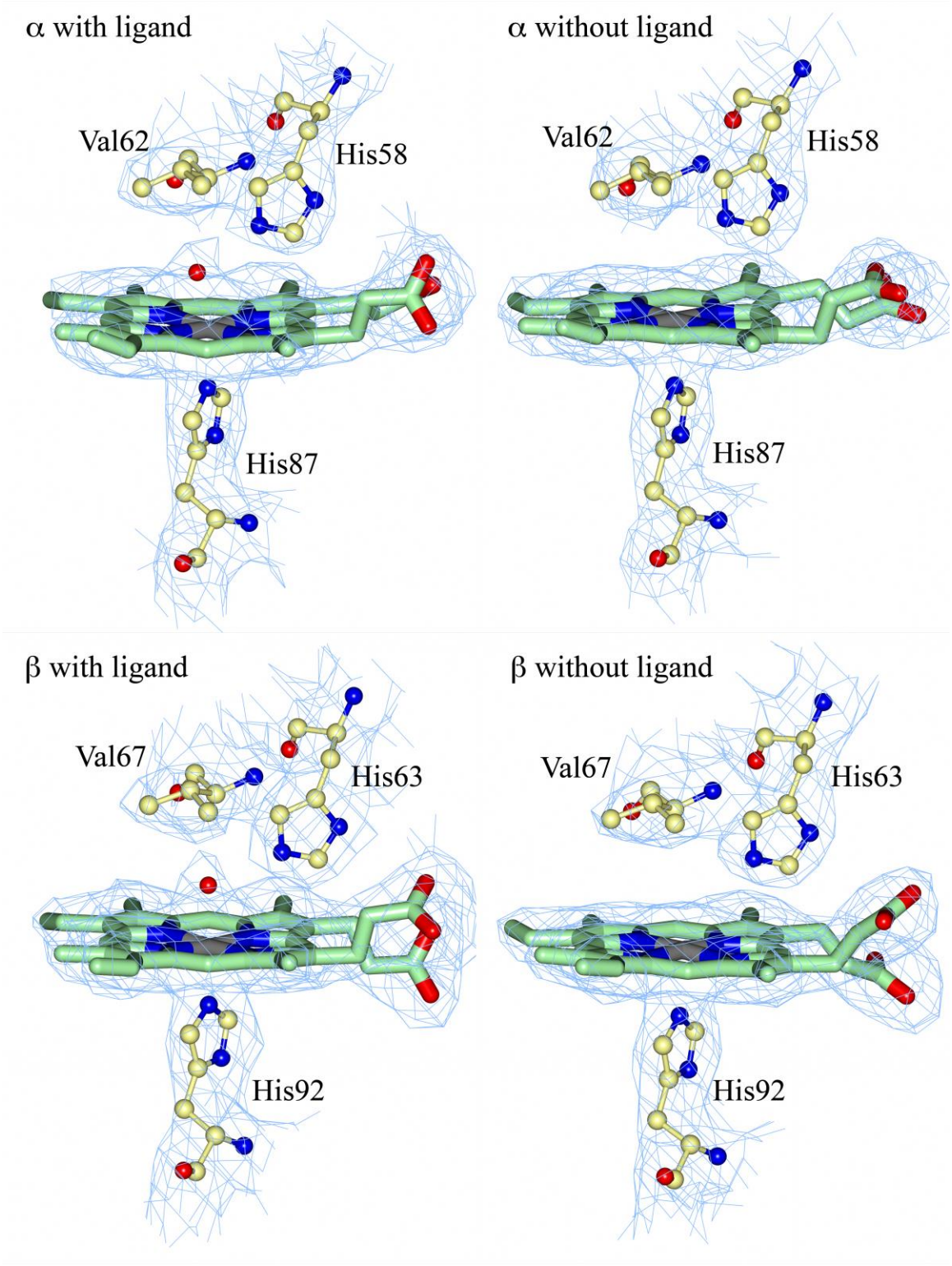

Figure 3.13 Electron density of the heme pocket in βGly83Cys demonstrating the ligand-bound and unbound forms of the aquamet model.

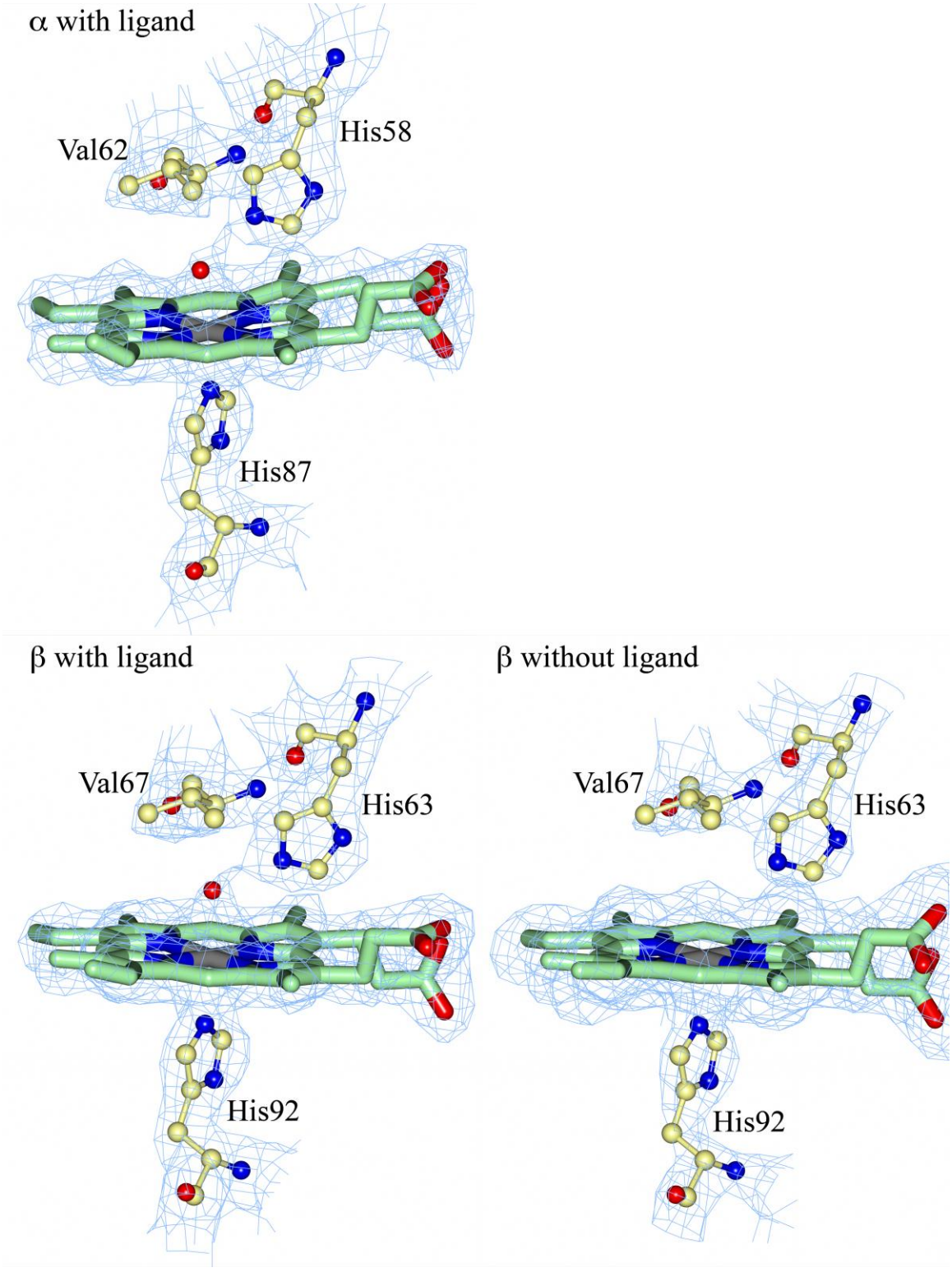

Figure 3.14 Electron density of the αAsn78Cys aquamet model with representative images of the α and β heme pockets.

Table 3.25 Distance from heme iron to various planes in the heme pocket. All numbers are in Ångstroms. Shown are distances from the iron to the proximal histidine (Prox His), distal histidine (Dist His), hydroxyl ligand (OH), the plane described by the 4 pyrrol nitrogens (Pyr N plane), and the plane described by the outer carbon atoms of the porphyrin ring (Por C plane).

	Prox His	Dist His	OH	Por N plane	Por C plane
αAsn78Cys α1	2.1	4.2	2.2	0.0	-0.2
αAsn78Cys α2	2.1	4.3	2.3	0.2	0.2
αAsn78Cys α3	2.1	4.3	2.2	0.1	0.2
αAsn78Cys α4	2.0	4.3	2.1	-0.1	-0.2
αAsn78Cys β1	2.2	4.2	2.3	0.0	-0.1
αAsn78Cys β2	2.2	4.2	---	0.1	0.1
αAsn78Cys β3	2.2	4.2	---	0.0	0.1
αAsn78Cys β4	2.1	4.3	---	-0.1	-0.2
βGly83Cys α1	2.2	4.2	---	-0.1	-0.2
βGly83Cys α2	2.3	4.2	---	0.0	0.1
βGly83Cys α3	2.2	4.2	---	0.1	0.2
βGly83Cys α4	2.3	4.2	---	0.0	-0.1
βGly83Cys α1’	2.3	4.1	---	0.0	-0.1
βGly83Cys α2’	2.1	4.3	---	0.1	0.2
βGly83Cys α3’	2.2	4.2	2.22	0.1	0.0
βGly83Cys α4’	2.2	4.2	---	-0.4	-0.2
βGly83Cys β1	2.3	4.1	---	-0.1	0.0
βGly83Cys β2	2.2	4.1	---	0.1	0.0
βGly83Cys β3	2.0	4.3	2.2	0.1	0.2
βGly83Cys β4	2.1	4.2	---	-0.1	-0.2
βGly83Cys β1’	2.1	4.3	---	0.0	-0.1
βGly83Cys β2’	2.1	4.2	---	0.0	0.1
βGly83Cys β3’	2.1	4.2	---	0.1	-0.1
βGly83Cys β4’	2.1	4.3	2.2	0.0	0.2
βGly83Cys-Phe41Tyr α1	2.1	4.2	2.3	0.0	0.0
βGly83Cys-Phe41Tyr α2	2.1	4.2	2.4	0.0	-0.1
βGly83Cys-Phe41Tyr α3	2.1	4.2	2.2	0.1	0.0
βGly83Cys-	2.1	4.2	2.3	0.0	0.1

Phe41Tyr α4					
βGly83Cys-Phe41Tyr α1’	2.1	4.2	2.5	0.0	0.0
βGly83Cys-Phe41Tyr α2’	2.1	4.1	2.2	-0.1	0.0
βGly83Cys-Phe41Tyr α3’	2.2	4.1	2.4	-0.1	0.0
βGly83Cys-Phe41Tyr α4’	2.1	4.2	2.4	0.1	0.0
βGly83Cys-Phe41Tyr β1	2.1	4.3	---	0.1	0.0
βGly83Cys-Phe41Tyr β2	2.1	4.3	2.3	0.0	0.1
βGly83Cys-Phe41Tyr β3	2.1	4.3	2.4	0.1	0.0
βGly83Cys-Phe41Tyr β4	2.1	4.3	2.2	0.0	0.1
βGly83Cys-Phe41Tyr β1’	2.2	4.3	2.4	0.1	0.0
βGly83Cys-Phe41Tyr β2’	2.1	4.3	2.4	0.0	0.0
βGly83Cys-Phe41Tyr β3’	2.1	4.3	2.3	0.0	-0.1
βGly83Cys-Phe41Tyr β4’	2.1	4.3	2.3	0.0	0.1

Table 3.26 Average heme pocket geometry in the oxidized octameric mutants.

	αAsn78Cys	αAsn78Cys	βGly83Cys	βGly83Cys	βGly83Cys-Phe41Tyr	βGly83Cys-Phe41Tyr
	α	β	α	β	α	β
Prox His	2.1 ±0.05	2.2 ±0.06	2.2 ±0.05	2.1 ±0.06	2.1 ±0.02	2.1 ±0.03
Dist His	4.3 ±0.02	4.2 ±0.07	4.2 ±0.06	4.2 ±0.07	4.2 ±0.03	4.3 ±0.03
OH	2.2 ±0.04	2.3	2.22	2.2 ±0.00	2.3 ±0.08	2.3 ±0.09
Pyr N plane	0.0 ±0.14	0.0 ±0.07	0.0 ±0.12	0.0 ±0.05	0.0 ±0.06	0.0 ±0.04
Pyr plane	0.0 ±0.27	0.0 ±0.17	0.0 ±0.17	0.1 ±0.12	0.0 ±0.01	0.0 ±0.07

In the βGly83Cys-Phe41Tyr mutant, the α-chain distal histidines are very similarly placed. The heme residues were superimposed on each other in order to get the best fit of the heme pocket for comparison. The distance between the nitrogens in the distal histidine is less than 0.2 Å, suggesting that the nitrogen placement is stabilized by the presence of hydroxyl in the heme pocket.

The β-chain distal histidines in the βGly83Cys mutant vary in placement in the heme pocket. Table 3.26 compares the distances between the iron and the proximal and distal histidines, as well as the displacement of the iron from the porphyrin ring plane. All mutants have the iron well in the plane of the porphyrin ring, regardless of whether the porphyrin nitrogens or the carbon pyrrole ring is used to create the plane (see Table 3.25 and Table 3.26). While little difference is seen among the mutants for iron positioning, the proximal and distal histidines vary in their positioning from the iron and the hydroxyl group (see Figure 3.12 to Figure 3.14 and Table 3.26).

The αAsn78Cys β heme pocket does not have a consistent conformation, suggesting both positive and negative forces in ligand binding. Half of the β chain heme pockets show a more closed conformation, where the distal histidine is closer to the iron and the aromatic residues around the heme are closer to the porphyrin. The other half adopts a conformation that is slightly distended outwards towards the outside of the tetramer. The heme pockets with the longer distance between the distal histidine and the iron would be more sterically favorable for ligand binding, whereas the shorter distance in the other monomers would inhibit ligand binding.

3.3.7 Effects of the βPhe41Tyr mutation

The βPhe41Tyr mutation has been shown to decrease the oxygen and carbon monoxide affinity of HbA. Ho and coworkers have shown that addition of the βPhe41Tyr mutation to the octameric βGly83Cys mutant also decreases the affinity of the octameric Hb. The βTyr41 adds additional hydroxyl group that has the capability to hydrogen bond and stabilize interactions that would decrease the oxygen affinity. The hydroxyl group of β_1Tyr41 hydrogen bonds with β_1Val98, β_1Asp99, β_1Asn102 and α_2Asp 94 (see Figure 3.15, Figure 3.16 and Table 3.27). Table 3.27 shows the hydrogen bonding distances between the βTyr41 hydroxyl and the other residues. Also note the proximity of βTyr41 to the heme pocket. Any changes that occur near this mutation, especially interactions with βVal98, would affect the accessibility and stability of the heme pocket. In HbA, the βPhe41 is stabilized by hydrophobic interactions, which are much weaker than hydrogen bonds. The weaker hydrophobic val der Waals interactions allow βPhe41 to move more freely than the βTyr41.

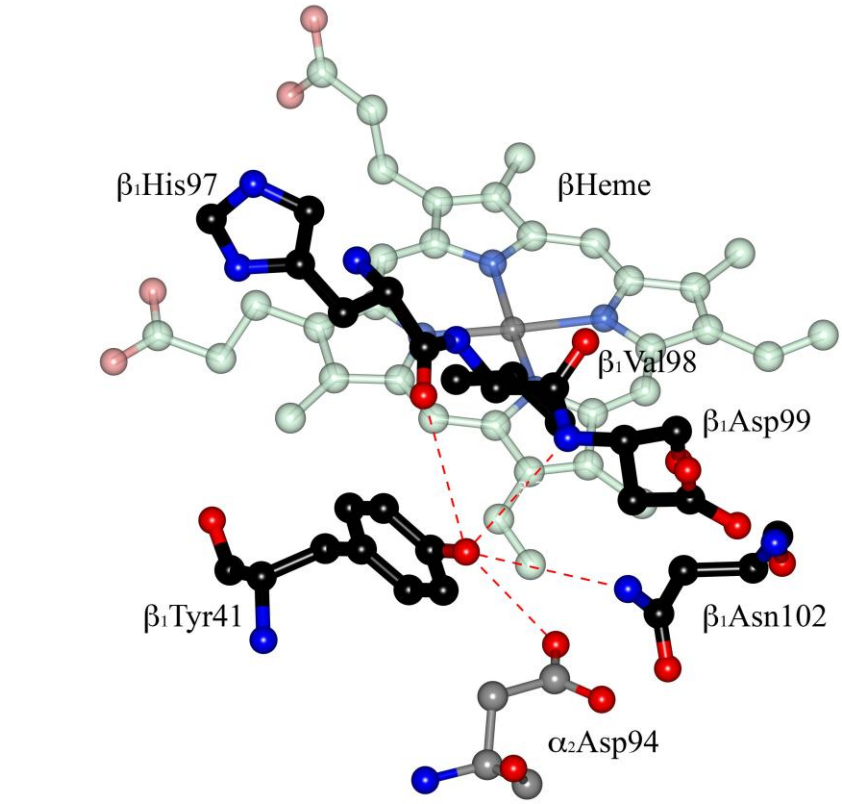

Figure 3.15 Hydrogen Bonding Interactions of the Mutated βTyr41 in βGly83Cys-Phe41Tyr

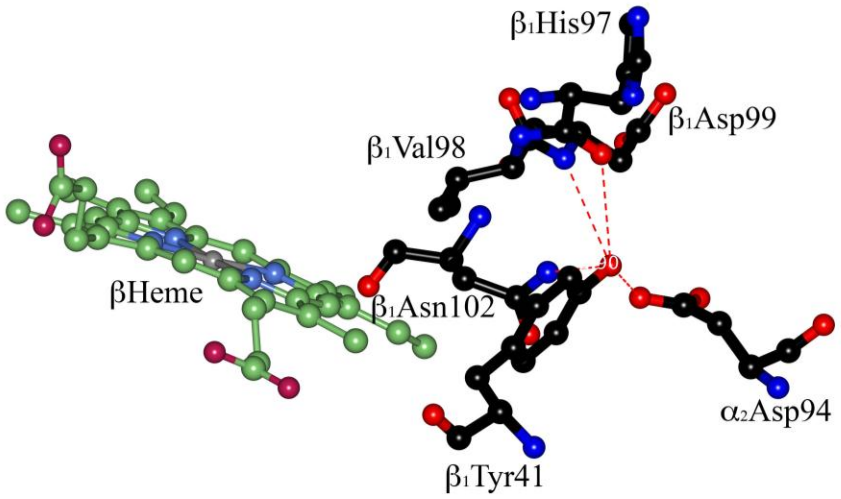

Figure 3.16 Proximity of βTyr41 and Interacting Residues to the β Heme

Table 3.27 Hydrogen bonding distances of mutant βPhe41Tyr interactions

Monomer	β_1Asp99 N	β_1His97 O	β_1Asn94 ND2	α_2Asp94 OD
β_1	3.3	3.7	3.9	2.9
β_2	3.4	3.7	3.9	3.0
β_1'	3.4	3.7	3.9	2.8
β_2'	3.3	3.7	3.9	2.9
β_3	3.3	3.7	4.0	2.8
β_4	3.4	3.7	3.9	3.0
β_3'	3.4	3.7	3.9	3.0
β_4'	3.4	3.7	3.9	3.1
Average	3.4±0.0	3.7±0.0	3.9±0.0	2.9±0.1

3.3.8 Cysteine positioning and activity

In the βGly83Cys-Phe41Tyr mutant, βCys93 has different positions, suggesting that it controls the activity of βGly83Cys-Phe41Tyr. Previous research has shown that trapping the βCys93 in the outward-facing state by binding to NEM or IAA changes the T-to-R transition and stabilizes the T state. The βCys93 does not adopt multiple conformations in the other mutants. The

βGly83Cys mutant contains the βCys93 pointing inwards, towards the core of the protein. αAsn78Cys mutant contains the outward-facing βCys93 conformation.

3.3.9 oHb vs. HbCN

The oxidized hemoglobin contains water molecules at many of the heme pockets. In an effort to obtain hemoglobin that is more like the oxygenated hemoglobin that would be used as blood substitutes, the octameric hemoglobin was purposefully oxidized and reacted with cyanide (CN^-). Cyanide is a two-atom ligand, like oxygen, where the carbon interacts with the porphyrin iron and the nitrogen, with its lone pair of electrons, can interact and hydrogen bond with the distal histidine, much like oxygen molecules in HbO_2.

Although the tetramers do not have a lot of change, the presence of CN^- appears to stabilize the disulfide bridge. In the oHb βGly83Cys, the large changes in the disulfide bridge geometry affected the relationship between the tetramers, causing the crystallization to be more variable. This is not seen in the cyanide-bound oHb, and allows the protein to form more stable interactions and crystallize in a higher space group. Figure 3.17 shows a superposition of the oxidized βGly83Cys and the cyanide-bound high salt βGly83Cys structure. The two tetramers are adopting slightly different conformations. The presence of cyanide appears to stabilize one particular conformation of the disulfide bridge, resulting in only one octamer conformation. This cyanide octamer conformation tilts the second tetramer 24° further than the oxidized form after superpositioning the tetramer. However, the presence of cyanide does not appear to affect the relative dimer position. Comparison of the tetramers with previously published structures shows that the cyanide-bound βGly83Cys are most similar to the PEG-crystallized structure of 1BBB (see Table 3.29 and Table 3.30).

The cyanide appears to bind only α monomers (see Table 3.28 and Figure 3.18). The distal histidine moves further away from the iron to accommodate the larger ligand. The presence of the disulfide bridge between the β monomers possibly destabilizes the cyanide binding in the β monomers. If the β monomers would need more flexibility at the ligand binding site due to constraints caused by β-β interactions, the average ligand binding to the β monomers may be similar to HbA but not consistant enough to see in crystallography. This may also be affecting the ability to crystallize the βGly83Cys-Phe41Tyr. If the cyanide introduces alot of quaternary changes in the β monomers that would then get distributed through the entire octamer it may be affecting the octameric structure of the βGly83Cys-Phe41Tyr and inhibiting crystallization. The small size of the hydroxyl in the oxidized mutants may not be enough to cause these disturbances in the octamer.

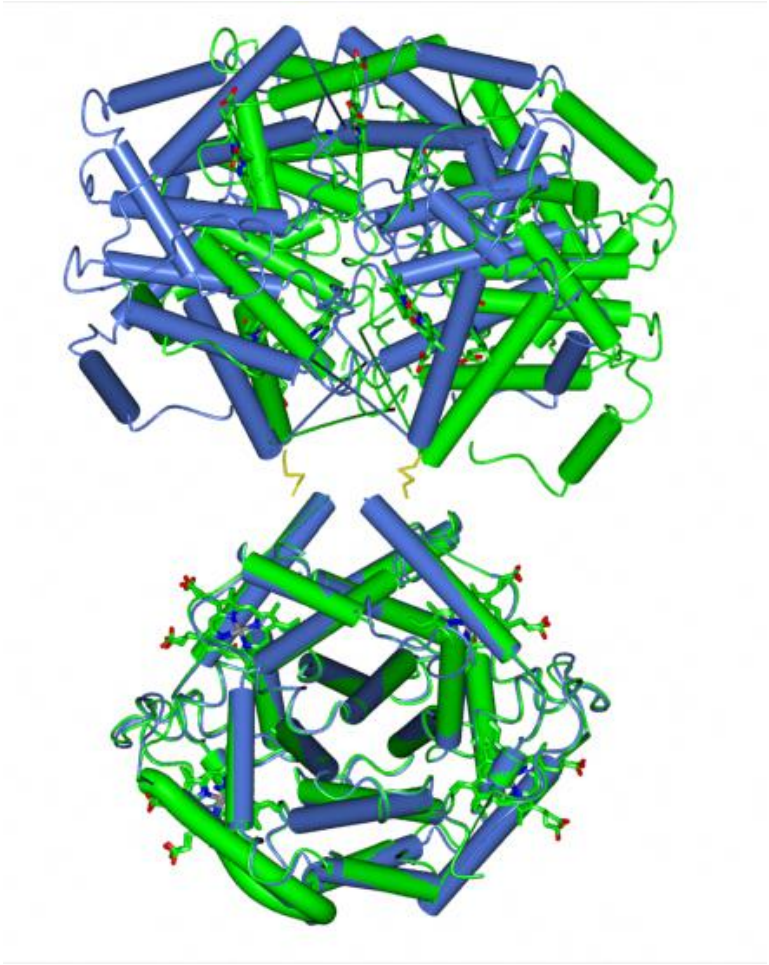

Figure 3.17 Superposition of one tetramer of the βGly83Cys CN (green) and oβGly83Cys (blue) shows the large difference in the position of the second tetramer due to changes in the disulfide bridge geometry. The second tetramer is shifted by 24°.

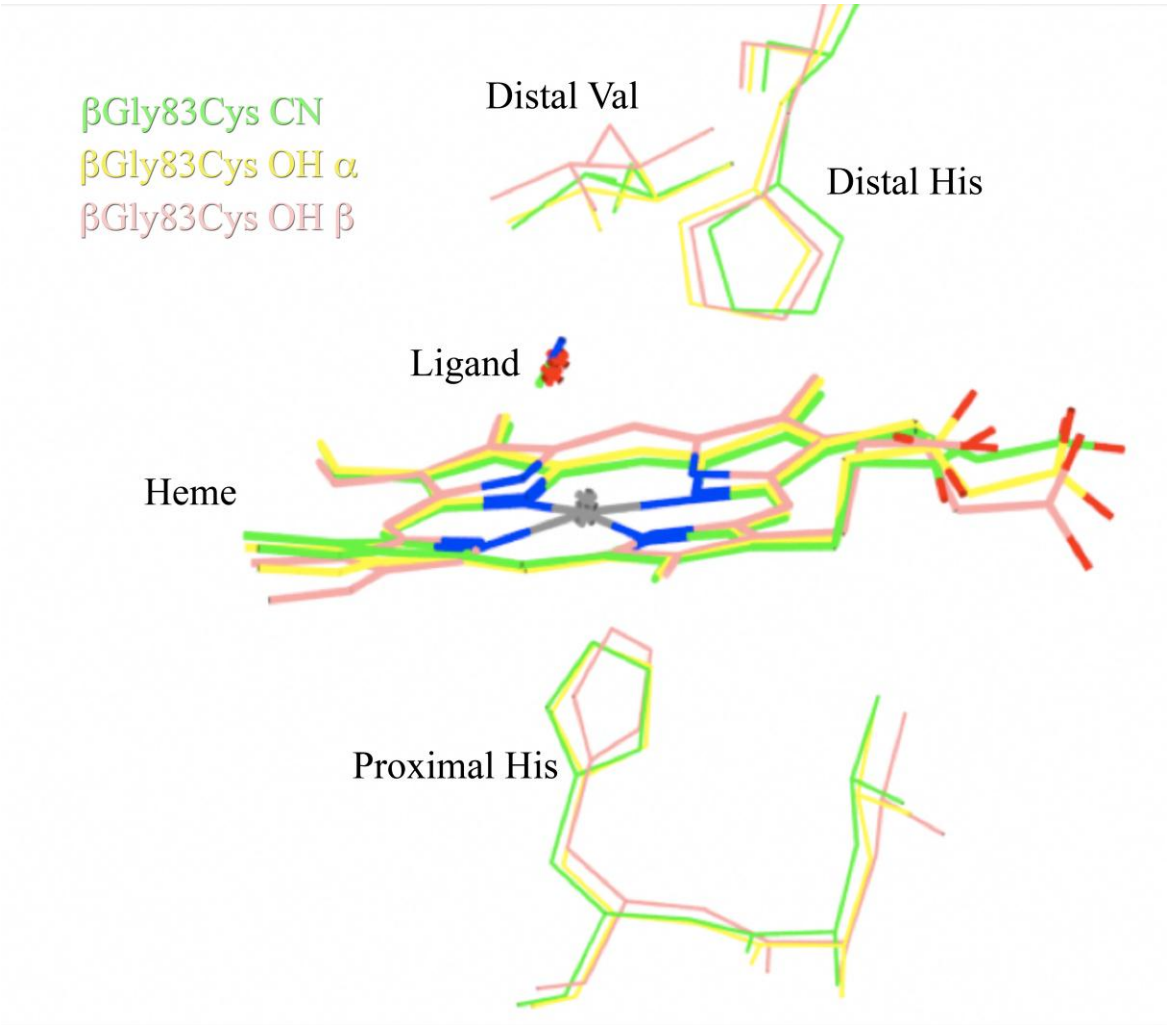

Figure 3.18 Cyanide-bound βGly83Cys heme compared to hydroxyl-bound βGly83Cys heme. The distal His moves away from the heme pocket to accommodate the cyanide.

Table 3.28 Heme group geometry of the high salt βGly83Cys CN structure. All numbers are in Ångstroms. Shown are distances from the iron to the proximal histidine (Prox His), distal histidine (Dist His), hydroxyl ligand (OH), the plane described by the 4 pyrrol nitrogens (Pyr N plane), and the plane described by the outer carbon atoms of the porphyrin ring (Por C plane).

	Por plane	Por N plane	Ligand	Proximal His	Distal His
with CYN					
α_1	-0.01	0.18	2.25	2.26	4.54
α_2	-0.03	0.10	2.25	2.38	4.80
α_1'	0.27	0.26	2.25	2.33	4.34
average	0.08±0.20	0.18±0.08	2.25±0.00	2.32±0.07	4.56±0.24
with OH					
α_2'	0.10	-0.09	2.50	2.25	4.23
β_1	-0.05	-0.12	2.51	2.26	4.11
β_2	0.06	0.00	2.58	2.36	4.23
β_2'	-0.04	0.05	2.55	2.26	4.31
average	-0.01±0.07	-0.35±0.08	2.55±0.037	2.29±0.07	4.22±0.11
no ligand					
β_1'	-0.01	-0.13	NA	2.26	4.29

Table 3.29 RMSD of dimers of the high salt octameric βGly83Cys CN and previously published structures (by chain) after α1β1 dimer superposition. Where only one dimer comparison is listed, the second dimer is related by the asymmetric unit and would give identical RMSD values.

	2DN3	1BBB $\alpha_1\beta_1$	1BBB $\alpha_2\beta_2$	1GZX $\alpha_1\beta_1$	1GZX $\alpha_2\beta_2$	1YZI	1A3N $\alpha_1\beta_1$	1A3N $\alpha_2\beta_2$	1MKO $\alpha_1\beta_1$	1MKO $\alpha_2\beta_2$	2DN1
$\alpha_1\beta_1$	0.93	0.34	0.31	1.19	1.17	1.16	1.23	1.25	0.71	0.53	0.97
$\alpha_2\beta_2$	0.89	0.40	0.32	1.13	1.11	1.15	1.17	1.19	0.70	0.50	0.92
$\alpha_1'\beta_1'$	0.91	0.38	0.34	1.16	1.14	1.14	1.20	1.22	0.71	0.54	0.93
$\alpha_2'\beta_2'$	0.92	0.37	0.30	1.18	1.17	1.17	1.22	1.23	0.73	0.52	0.96
	0.91 ±0.02	0.37 ±0.03	0.32 ±0.02	1.16 ±0.01	1.15 ±0.03	1.16 ±0.01	1.20 ±0.03	1.22 ±0.03	0.72 ±0.01	0.52 0.02	0.95 ±0.02
		0.34±0.03		1.16±0.02			1.21±0.03		0.62±0.11		

In order to compare the cyanide-bound βGly83Cys with the previously published HbA structures, the RMSD of the dimers was compared in the same way as the oxidized mutants (see Table 3.29 and Table 3.30). The $\alpha_1\beta_1$ dimer was superimposed and the RMSD obtained (column $\alpha_1\beta_1$), then the RMSD for $\alpha_2\beta_2$ was obtained (column $\alpha_2\beta_2$). The angle denotes the rotation angle necessary to superimpose the vector between the first and last Cα of each $\alpha_2\beta_2$ dimer. Both the $\alpha_1\beta_1$->$\alpha_2\beta_2$ rotation and the $\alpha_2\beta_2$->$\alpha_1\beta_1$ rotation were calculated because intrinsic differences in the RMSD and superpositions of each dimer would affect the RMSD and angle rotation calculated for the second dimer. The first two letters (i.e. in $\alpha_1\beta_1$->$\alpha_2\beta_2$ the $\alpha_1\beta_1$) indicate the chains that were used as the $\alpha_1\beta_1$ dimer and the second two letters (i.e. in $\alpha_1\beta_1$->$\alpha_2\beta_2$ the $\alpha_2\beta_2$) indicate the chains that were used as the $\alpha_2\beta_2$ dimer. As in the oxidized mutants, βGly83Cys CN is closest to the structure 1BBB, which was crystallized in PEG. The fact that all of the mutants are closest to the 1BBB structure, regardless of the crystallization conditions, demonstrates the propensity of the disulfide bridge to restrict the conformations that are available for the octameric hemoglobin. The crystallization conditions of the high salt structure are closest to the crystallization conditions of the original R structure, so the expected conformation of the high salt βGly83Cys would be closer to the R structure.

Table 3.30 Comparison of high salt βGly83Cys with previously published structures. See text for detailed methods.

1A3N			
	$\alpha_1\beta_1$ (Å)	$\alpha_2\beta_2$ (Å)	Angle (°)
$\alpha_1\beta_1$ -> $\alpha_2\beta_2$	1.23	8.62	29.8
$\alpha_2\beta_2$ -> $\alpha_1\beta_1$	1.17	8.55	29.9
$\alpha_1'\beta_1'$ -> $\alpha_2'\beta_2'$	1.20	8.53	30.3
$\alpha_2'\beta_2'$ -> $\alpha_1'\beta_1'$	1.22	8.45	29.9
Average	1.20±0.03	8.54±0.07	30.0±0.2
1GZX			
	$\alpha_1\beta_1$ (Å)	$\alpha_2\beta_2$ (Å)	Angle (°)
$\alpha_1\beta_1$ -> $\alpha_2\beta_2$	1.19	9.27	31.1
$\alpha_2\beta_2$ -> $\alpha_1\beta_1$	1.13	9.18	31.2
$\alpha_1'\beta_1'$ -> $\alpha_2'\beta_2'$	1.16	9.15	31.6
$\alpha_2'\beta_2'$ -> $\alpha_1'\beta_1'$	1.18	9.08	31.3
average	1.16±0.02	9.17±0.07	31.3±0.2
2DN1			
	$\alpha_1\beta_1$ (Å)	$\alpha_2\beta_2$ (Å)	Angle (°)
$\alpha_1\beta_1$ -> $\alpha_2\beta_2$	0.97	4.17	10.1
$\alpha_2\beta_2$ -> $\alpha_1\beta_1$	0.92	4.04	10.1
$\alpha_1'\beta_1'$ -> $\alpha_2'\beta_2'$	0.93	4.00	10.5
$\alpha_2'\beta_2'$ -> $\alpha_1'\beta_1'$	0.96	3.96	10.1
average	0.95±0.02	4.04±0.08	10.2±0.2
2DN3			
	$\alpha_1\beta_1$ (Å)	$\alpha_2\beta_2$ (Å)	angle (°)
$\alpha_1\beta_1$ -> $\alpha_2\beta_2$	0.93	3.88	9.5
$\alpha_2\beta_2$ -> $\alpha_1\beta_1$	0.89	3.76	9.5
$\alpha_1'\beta_1'$ -> $\alpha_2'\beta_2'$	0.91	3.72	9.8
$\alpha_2'\beta_2'$ -> $\alpha_1'\beta_1'$	0.92	3.69	9.4
average	0.91±0.02	3.76±0.08	9.5±0.2
1MKO			
	$\alpha_1\beta_1$ (Å)	$\alpha_2\beta_2$ (Å)	Angle (°)
$\alpha_1\beta_1$ -> $\alpha_2\beta_2$	0.71	2.10	6.4
$\alpha_2\beta_2$ -> $\alpha_1\beta_1$	0.70	1.93	6.4
$\alpha_1'\beta_1'$ -> $\alpha_2'\beta_2'$	0.71	1.94	6.7
$\alpha_2'\beta_2'$ -> $\alpha_1'\beta_1'$	0.73	1.98	6.3
average	0.72±0.01	1.99±0.08	6.5±0.2
1BBB			
	$\alpha_1\beta_1$ (Å)	$\alpha_2\beta_2$ (Å)	Angle (°)
$\alpha_1\beta_1$ -> $\alpha_2\beta_2$	0.31	0.73	1.8
$\alpha_2\beta_2$ -> $\alpha_1\beta_1$	0.40	0.83	1.7

$\alpha_1'\beta_1'$ -> $\alpha_2'\beta_2'$	0.38	0.69	1.3
$\alpha_2'\beta_2'$ -> $\alpha_1'\beta_1'$	0.37	0.70	1.7
average	0.37±0.03	0.74±0.06	1.6±0.2
1YZI			
	$\alpha_1\beta_1$ (Å)	$\alpha_2\beta_2$ (Å)	Angle (°)
$\alpha_1\beta_1$ -> $\alpha_2\beta_2$	1.16	5.58	18.6
$\alpha_2\beta_2$ -> $\alpha_1\beta_1$	1.15	5.44	18.7
$\alpha_1'\beta_1'$ -> $\alpha_2'\beta_2'$	1.14	5.44	18.9
$\alpha_2'\beta_2'$ -> $\alpha_1'\beta_1'$	1.17	5.45	18.5
average	1.16±0.01	5.48±0.07	18.7±0.2

3.4 DISCUSSION

While currently a search for human hemoglobin models in the PDB will bring up over 200 entries (214 as of March 2012), none of the entries has been solved under crystallization conditions that have such a high pH value. Previous researchers have looked for patterns in crystallization and pH, and have supported a conclusion that the overall surface charge of the crystallized protein must be close to 0 in order for crystallization to be successful. This makes sense because protein-protein contacts must occur in crystallization, where typically positively-charged amino acids interact with typically negatively-charged amino acids. Tetrameric hemoglobin has a pI (7.89 according to Protparam [98]) higher than physiological pH (7.2-7.4). HbA contains 44 lysines, the majority of which protrude into the solution. In the tetrameric form, these lysines can interact with a comparative portion of the hemoglobin tetramer. However, when restricted by the addition of disulfide bridges across two tetramers, it forces the positive areas of hemoglobin to interact with other lysine-containing regions. This necessitates

crystallization at a higher pH, where the lysines would be partially protonated. This also may explain the structural stabilization of particular R states in tetrameric hemoglobin.

The presence of two disulfide bridges per octamer allows temporary reduction of either of the two disulfide bridges without dissociation of the entire complex. If one disulfide bridge becomes reduced, the proximity and orientation of the two tetramers allows the disulfide bridge to reform at a rate higher than would be expected if only one covalent interaction stabilized the octamer. The octamerization allows for some amount of "give" in the dimers without a total dissociation.

3.4.1 Effects of the βPhe41Tyr mutation

The βPhe41Tyr mutation has been shown to decrease oxygen binding in the single mutant, however the mechanism was unknown because a crystal structure was nonexistent for this particular mutant. In HbA, βPhe41 is part of the hydrophobic pocket that accepts the heme group in the β subunit. When superimposing the $\alpha_1\beta_1$ dimer of 1A3N, the T state tetramer, and 2DN3, the carbonmonoxy tetramer, the Cα atom of βPhe41 moves 0.8 Å. In the βGly83Cys-Phe41Tyr mutant, the βTyr41 instead interacts with backbone atoms from the β subunit, stabilizing the structure. This stabilization affects the oxygen binding properties but does not seem to affect the disulfide bridge, as the bridge still has a significant amount of motion as the single βGly83Cys mutant does. The interactions with β_1Asp99 backbone, β_1His97 backbone, α_2Asp94, and β_1Asn102 stabilize the R state conformation. If the βPhe41 is mutated and modeled as a Tyr in the PDB 1A3N, the hypothesis can be made that the T state and lower affinity is stabilized by interactions between β_1Tyr41 and β_1Asp99 backbone, β_1His97 backbone, β_1Asn102. Note that the shift that occurs to bring βTyr41 closer to the α_2 subunit comes at the

expense of the hydrogen bonds between βTyr41 and the interacting residues on the β_1 subunit (see Figure 3.15 and Figure 3.16).

3.4.2 Effects of cyanide on crystallization

One of the problems experienced in crystallization of rHb is the movement possible from the disulfide bridge. While this could be independent of the ligand-bounding state, the fact that the rHb crystallized in a higher space group in the presence of CN suggests that the ligand binding stabilizes a specific octameric conformation. Interestingly, CN only bound to the α monomers, which has been shown to bind more tightly to carbon monoxide and oxygen.

3.4.3 Comparison to HbA

The octameric hemoglobin tetramers are closest in quaternary structure to the R2 ligand-bound state. Comparison shows similarities in ligand-binding state and crystallization conditions. The crystal structure closest to oHb is 1BBB. oHb and 1BBB crystals were grown in PEG conditions, which has been shown to affect the adopted R state. However, NMR research by Ho and coworkers has shown that these R state models are adopted by HbA in solution, and that crystallized models show individual conformational states that are part of the R ensemble.

One concern is the fact that these models were determined at a much higher pH than physiological pH and much higher than the normal 6.5-7.5 pH range that is usually seen in crystallized models. The higher pH brings into question how much of the models are affected by crystallization conditions. The similarities of the overall structure of octameric hemoglobin in the presence and absence of cyanide suggests that the crystallization conditions are not affecting

the model any more than previously determined HbA, but that the ligand binding state has a much greater effect on the adopted structure.

Even when oxidized, the αAsn78Cys model shows structural aspects of the T state. The models of the αAsn78Cys do not show βCys93, directly adjacent to the proximal histidine βHis92, to be in the same conformation in all monomers. In half of the β chains, βCys93 is pointing inwards, whereas in the other half of the model, βCys93 is pointing towards the outside of the model. Previous research has shown that the T-R transition is accompanied by a flipping of βCys93 from the core to the surface of the Hb. Indeed, Hb can be forced to keep the R conformation by linking the βCys93 with large sulfide linkers that keep βCys93 pointing towards the solvent. In the T state, βCys93 is sequestered and does not react easily with cysteine-altering compounds such as IAA and NEM. However, once hemoglobin is in the R state, the βCys93 is readily available for reaction with cysteine-altering compounds. Hemoglobin can be stabilized in the R state by reacting βCys93 with IAA prior to stripping the heme groups of ligands. The changes suggest that ligand affinity is modulated by restraining the βCys93 position, affecting the heme environment.

3.4.4 Summary and Future Directions

The disulfide bridges stabilize an octameric hemoglobin and prevent dissociation of the cooperative tetramer into dimers. The disulfide bridge in the βGly83Cys and βGly83Cys-Phe41Tyr mutants can adopt a wider range of geometry than the disulfide bridges in the αAsn78Cys, which are more restricted in part due to the steric interactions between the two linked tetramers. The disulfide bridge can also be stabilized by the presence of cyanide,

suggesting that the oxygen-bound octameric hemoglobin would also have less relative motion across the disulfide bridge. The presence or absence of ligand in the heme pockets, specifically the α heme pockets, have a significant effect on the overall octameric structure.

The presence of cyanide appears to stabilize the disulfide bridge conformations by stabilizing the α heme pockets and affecting the rest of the model through αβ interactions. The β heme pockets do not appear to contain cyanide, suggesting that the disulfides destabilize the tetramer.

4.0 RAP1B

4.1 INTRODUCTION

Rap1b is a protein that is phosphorylated on Ser179 in order to activate it. Previous attempts at crystallization have shown the core structure to be an alpha-beta protein [87-89]. Rap1 is an integral protein in cell-cell adhesion [162]. In order to identify the regulation of this protein, we must examine the full-length Rap. However, all of the previous models were terminated between residues 166-168, eliminating the portion of the protein crucial for regulation by phosphorylation and isoprenyl group addition [80, 81]. In the Rap proteins, Rap1A and Rap1B differ only in nine residues in the C terminal tail. This gives them a 95% sequence identity. Both proteins contain a phosphorylated serine, Ser180 in Rap1A and Ser179 in Rap1B. Because the first 171 residues of Rap1A and Rap1B are identical, the determined structures of Rap1 in complex with other proteins have been determined with little analysis of the conformations of the C terminus.

A portion of Rap1 between residues 32-44 has been shown to be necessary for interaction with Ras and GAP [163]. Subsequent mutations of Gly12Val and Ala59Thr were shown to increase its "reversion-inducing activity" [164]. Cys181 was mutated to Gly to remove the isoprenylation site. Two separate mutations were added to the double-mutant Gly12Val-Cys181Gly: Ser179Asp and Ser179Ala. The Ser179Asp mutation mimics the phosphorylation of Ser179 by adding a negative charge to residue 179. Aspartic acid was used instead of

glutamic acid because the aspartic acid contains a distance and geometrical restriction that is similar to a phosphorylated serine. Using a glutamic acid instead of an aspartic acid would extend the distance between the Cα and the negative charge and add a second area of rotation between the Cβ and Cγ atoms that would not be present in the phosphorylated serine. The other mutation is a Ser179Ala mutation, which would eliminate whatever interactions exist from the presence of the normal Ser179.

4.2 MATERIALS AND METHODS

4.2.1 Crystallization

Protein was supplied by Dr. Daniel Altschuler, containing GTPγS as the GTP analogue. Hanging drop vapor diffusion crystallizations were set up by Dr. Shoucheng Du. Crystals were obtained from crystallization conditions 0.3 M magnesium chloride, 0.1 M bicine pH 9.0, 25% w/v PEG 2000, 15% w/v glycerol and 0.2 M di-potassium phosphate, 20% w/v PEG 3350 for the 29A and 29D proteins, respectively. 22% ethylene glycerol was added to the mother liquor for cryoprotection of the 29D crystals.

4.2.2 Data Collection and Processing

Data collection was done at the SER-CAT BM line at the Advanced Photon Source at Argonne National Laboratory for the 29A and on a Saturn 944 CCD with a Rigaku FRE Superbright generator for the 29D data sets by Dr. Shoucheng Du and Dr. Unmesh Chinte. Data was

processed by Dr. Unmesh Chinte to 2.1 Å and 2.2 Å for the 29A and 29D, respectively, using HKL2000 [103].

4.2.3 Model Refinement

Initial models contained the core of the protein and the mtz data files, which were a product of Dr. Unmesh Chinte's refinement. The refinement statistics for the received models and final models are shown in Table 4.1. The C-terminus was determined by determining the electron density that reappeared after refinement when the terminus position was varied. All of the rotamers were checked against the electron density and standard rotamer conformations. The models were checked using simulated annealing omit maps to determine whether GTP or GDP was present in the active site. Waters and ions were added prior to checking the C-terminal placement. The placement of the C-terminus was modified to better fit the electron density. Final rounds of refinement were done in PHENIX [107]. Final model statistics can be found in Table 4.1.

Table 4.1 Data Refinement Statistics for Rap Ser179D and Rap Ser179A.

	Rap Ser179Asp Before	Rap Ser179Asp After	Rap Ser179Ala Before	Rap Ser179Ala After
Resolution (Å)	2.2	2.2	2.1	2.1
R_{work} (%)	24.83	17.86	24.72	18.69
R_{free} (%)	29.83	23.36	29.53	23.04
Bond angle RMSD	2.02	1.08	1.95	1.03
Bond length RMSD	0.021	0.008	0.021	0.008
Ramachandran statistics				
Most favored	89.9	92.8	88.5	96.0
Allowed	4.0	5.1	4.9	3.0
Disallowed	6.1	2.1	6.6	1.0
Clashscore	60.31	17.52	34.41	11.59
Rotamer outliers	10.7	3.0	8.6	2.3
Cβ outliers	2	1	0	0

4.3 RESULTS

4.3.1 Overall Structure and Changes in the C-terminus

Table 4.1 shows the refinement statistics of the two Rap mutants, listed as Ser179Asp (Gly12Val-Ser179Asp-Cys181Gly) and Ser179Ala (Gly12Val-Ser179Ala-Cys181Gly). The overall core fold is the same as previous Rap1 and Ras models. The Gly12Val mutation was to keep the Rap1B in an active conformation, and the Cys181Gly mutation removed the isoprenylation site. Both structures contain GDP instead of the GTPγS expected in the

nucleotide binding site. Both Rap models have the same core structure as previous cocrystallized Rap models (see Figure 4.1). The C-terminus is away from the core structure, so that it is not stabilized by interactions with the core of the protein. This may allow the core to provide a stable site to interact with the effectors and modulating proteins, while allowing the terminus to interact with and affect the activity of multiple different proteins. The structural flexibility of the loop would allow it to interact less tightly to a greater variety of proteins rather than interact with and affect one or two proteins very strongly. Considering that Rap must activate multiple proteins, the flexibility of the C-terminal loop supports the ability to interact with a variety of proteins.

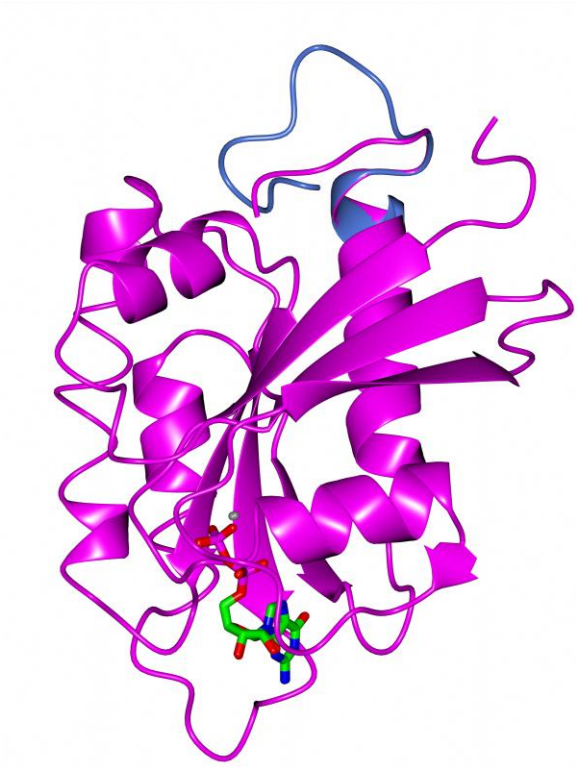

Figure 4.1 Superposition of Ser179Ala (magenta) and Ser179D (blue) showing the differences in the C-terminal loop position

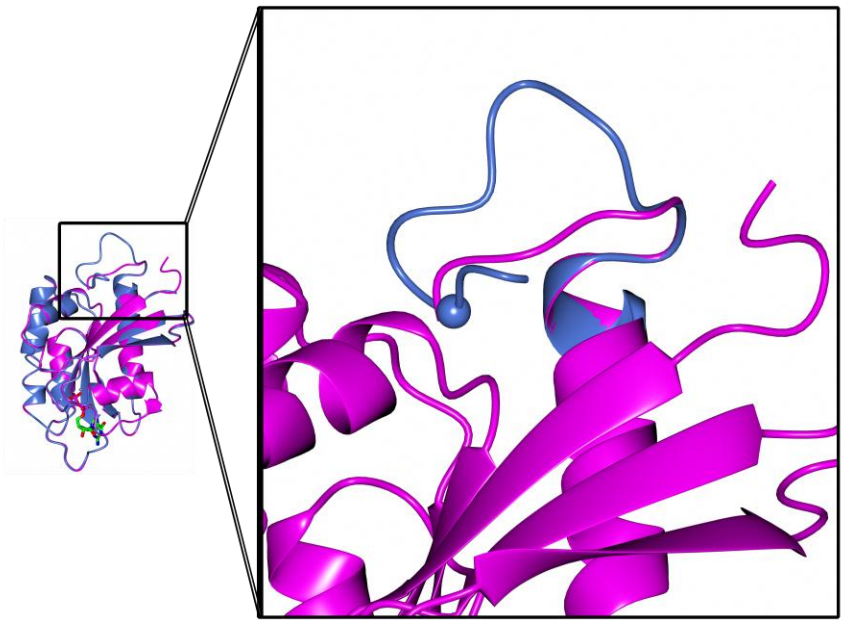

Figure 4.2 C-terminal region showing the Ser179Asp mutant (blue) with the mutation point marked with a ball at the Cα and Ser179Ala mutant (magenta). The C-terminal loop extends in the Ser179 mutant and is stabilized by interactions between the negatively-charged Asp179 and the protein core.

4.3.2 GTP site

Both Rap1B structures contain GDP and Mg^{2+} in the active site (see Figure 4.2 and Figure 4.3), suggesting that this model is in an inactive form. GTP hydrolysis occurred at some point between obtaining the protein and collecting crystal data. Rap1B is regulated by both the presence of GTP vs. GDP, and the phosphorylation at Ser179. The Ser179Asp mutant contains more interactions and more stable interactions in the C-terminal region compared to the Ser179Ala mutant.

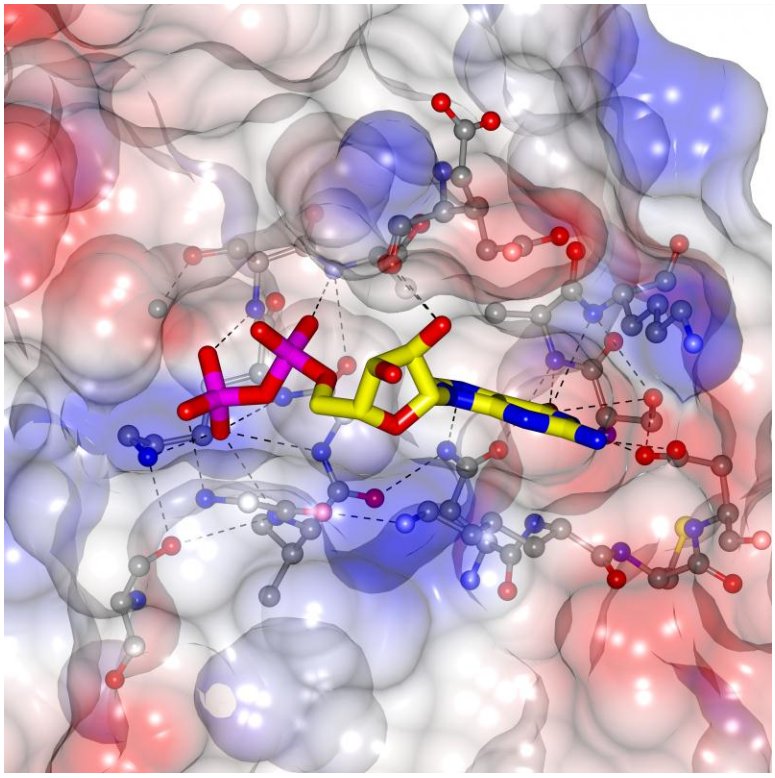

Figure 4.3 GDP binding site in Rap Ser179Ala. The binding site and interactions are identical in the Ser179Ala and Ser179Asp mutants. Dotted lines show the hydrogen bonding interactions between the protein and GDP.

The GDP is stabilized in position with interactions very similar to the interactions seen in the GTP analogue structures. The guanine base contains hydrogen bonds or charged interactions with the backbone nitrogens of residues Lys117, Ala148 and Lys149, and forms a network of hydrogen bonds or charged interactions with the sidechains of Asp119, Ser147 and Asn116 (see Figure 4.3 and Figure 4.4). One water molecule is stabilized close to the outer surface of the protein. The ribose sugar uses a much smaller network of hydrogen bonds and charged interactions with the backbone of Glu30 and Gly13, and interacts with the NZ atom of Lys117. Whether Lys117 stabilizes the sugar, or the significant interactions elsewhere between the GDP and Rap1B stabilize the specific conformation of Lys117 is unclear. The phosphates create a significant web, with each oxygen atom interacting with multiple residues in Rap1B. The

backbone atoms of Tyr32, Gly15, Gly13 and Ala18 cradle the α phosphate. The α phosphate also interacts with 4 water atoms. The β phosphate interacts with the same backbone atoms of Gly13 and Gly15, while creating new contacts with the backbone of Val14, Ser11, Lys16 and Ser17. The sidechains of Lys16 and Ser17 also interact with the β phosphate. The interactions stabilize the Mg^{2+} atom that is necessary for GTPase activity, along with four water atoms, two of which coordinate with the Mg^{2+} atom.

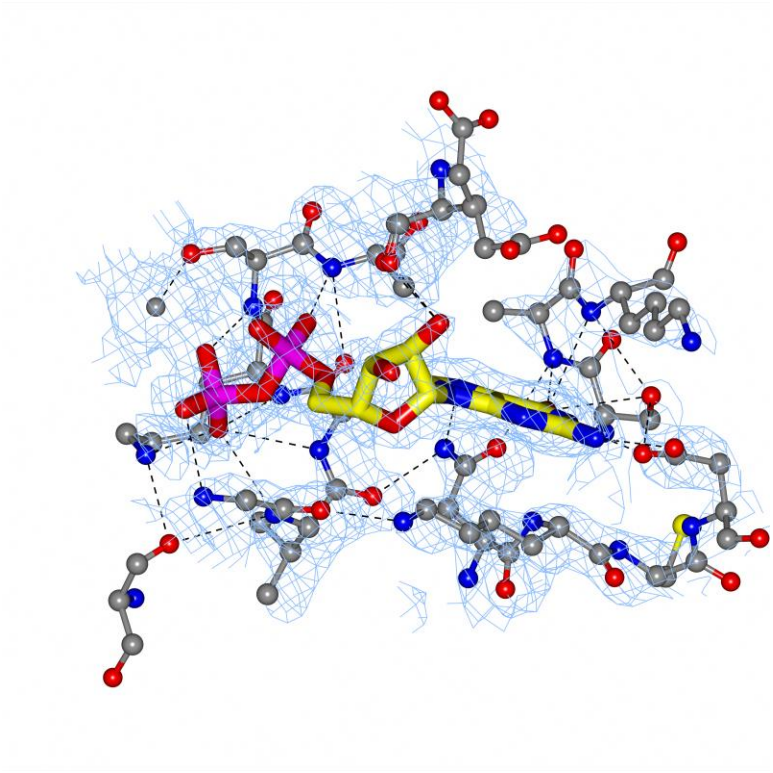

Figure 4.4 Electron density of the GDP and surrounding residues in Rap1B Ser179A. The Rap1B Ser179D contains the same contacts.

Attempts to place GTP in the nucleotide binding site support the conclusion that these mutants contain GDP. Positioning GTP in the binding site electron density distorts the geometry of the gamma phosphate of GTP. It would be necessary to replace the Mg^{2+} with the gamma phosphate, eliminating the magnesium. The current magnesium position is supported by other

Ras and Rap1 models, all of which have a magnesium ion in a comparable placement to the current models.

4.3.3 Protein changes close to the active site

One significant alteration in the refined Rap1B structures that has not been seen in previous Rap models in the positioning of Thr35. Previous Rap1 models show Thr35 pointing towards the active site of Rap where GTP binds [87-89]. Thr35 points away from the GDP binding site in our models. This area does not interact with symmetry-related Rap monomers, suggesting that the alteration is not due to crystal contacts.

4.3.4 Changes in the Switch II Region

Although the 29A and 29D models are very similar, with a superimposed Cα RMSD of 1.2 Å, there are two main regions that differ. One region is the C-terminal tail, discussed above. The other region is between Thr61 and Thr65 (see Figure 4.5). The electron density in this switch II region is not as well-defined as the rest of the model, with the exception of the C terminus. The switch II region is further away from the core of the protein in the Ser179A mutant. The Ser179D mutant contains the switch II region closer to the core, in a more stable conformation. The electron density for the Ser179D mutant is more visible in this region. Attempts to refine the switch II region of the Ser179A mutant in the conformation that the Ser179D mutant adopts do not improve the electron density and increase the R_{work} and R_{free} values.

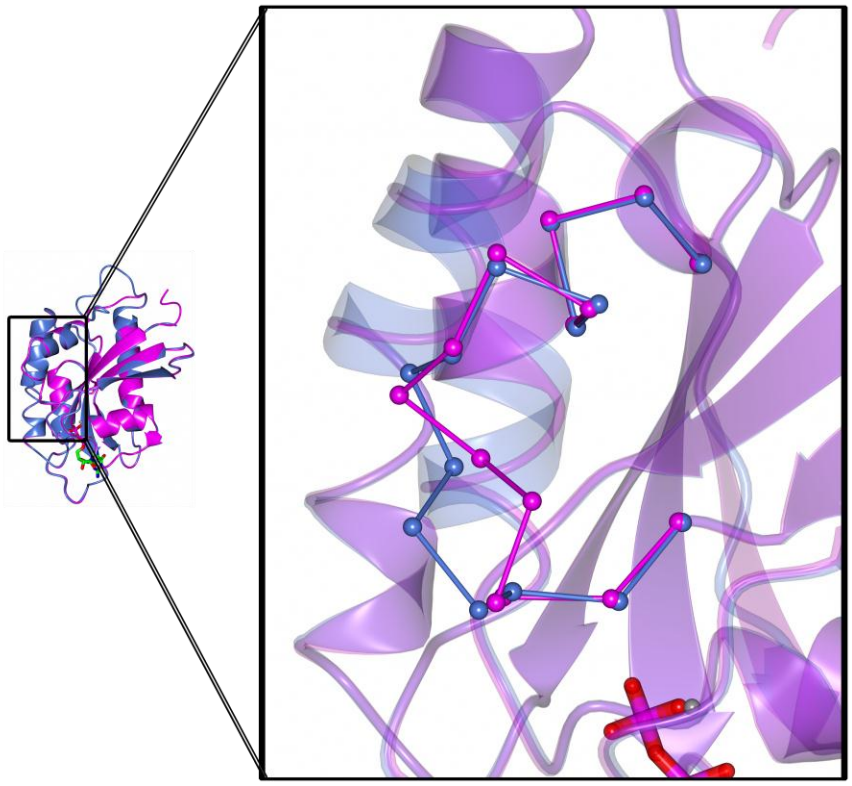

Figure 4.5 The Cα trace of the Ser179Asp mutant (blue) and Ser179Ala mutant (magenta).

4.4 DISCUSSION

We cannot rule out changes that occur due to posttranslational modifications of Rap1B. Previous work has shown that the lipid modification of Rap1B is necessary for stimulating C-Raf activity through Rap1B [82]. Both mutant models demonstrate active mutants that are the first Rap1 models to show the shift in Thr35 that occurs with GDP binding.

It is possible that in the absence of a "partner protein" Rap1B hydrolyzes GTP at a much faster rate than when in complex. The crystals were obtained 3-5 months after crystallization setup (Dr. Shoucheng Du, personal communication). However, the Gly12Val mutation should

have kept Rap1B in its active state and would have stayed complexed with GTP much longer than the WT Rap1B.

4.4.1 Position of Thr35 as a significant modulator

The conformation of Thr35 significantly changes when in the presence of GDP. The GTP analogues present in other models from the Ras family show that the equivalent of Thr35 interacts with the gamma phosphate of the GTP analogue. The Rap1 models in complex with other proteins all demonstrate that Thr35 interacts with the Mg^{2+} and gamma phosphate of the GTP analogue. The rotation of Thr35 away from the GTP binding site in the presence of GDP suggests that the structural change is necessary for GDP-GTP exchange. A guanine exchange factor (GEF) interacts with the Rap1 in order to complete the conversion from the inactive GDP-bound state to the active GTP-bound state [165].

4.4.2 Summary

These are the first structures to show the changes that occur to Rap1B in the presence of GDP. The models also demonstrate that the presence of the negative charge at amino acid position 179 significantly changes the placement of the C terminus, creating different charged interactions in the Rap protein. The alterations that occur due to the different charges suggest two different switches in the interactions between Rap and its activated proteins. One switch is the active/inactive switch previously documented that is triggered by the presence of GTP or GDP. The secondary switch is the charged interactions that occur during phosphorylation. The change

in the C terminal position changes the interactions between Rap and other proteins, allowing for a level of control that is not present in the GTP/GDP switch.

5.0 CONCLUSION

Multiple systems feed into the cellular respiration pathways, including hemoglobin, glycerol metabolism, and G-proteins such as Rap1B. I have determined structural models of *Enterococcus casseliflavus* glycerol kinase mutants with glycerol or ethylene glycol, *Pseudomonas aeruginosa* glycerol kinase with glycerol or ethylene glycol, octameric human hemoglobin mutants, and two Rap1B mutants. The structures demonstrate that small changes in one region of the model can significantly affect the protein activity and adopted conformation of other portions of the model.

The structures of the *En. casseliflavus* GK mutants allow us to determine how the mutations activate the enzyme, and infer how the phosphorylation of His232 activates *En. casseliflavus* GK. Although the His232Glu mutant was designed to mimic the negative charge caused by histidine phosphorylation, the distance between the charge and the Cα is smaller in the His232Glu than it would be in the His-PO4. The shorter distance does not allow Glu232 to stabilize the activation loop. However, the His232Arg mutant contains a longer distance between the Cα and the positive charge on Arg232 than the His232Glu mutant. The His232Arg mutant contains contacts between the activation loop and Domain I that stabilize a conformation that causes closure of the glycerol binding cleft. The presence of loop interactions creates a domino effect that closes the active cleft and orients the substrates properly for glycerol phosphorylation. The His232Glu mutant is not positioned correctly to make the interactions seen

in the His232-PO_4 activated form because it is not long enough. The same loop interactions are seen in the ethylene glycol-bound *En. casseliflavus* GK, which shows the monomer rotation necessary to activate GK once glycerol binds. This suggests that factors that disrupt the activation loop would decrease the activity of GK in Gram positive bacteria. Factors that would stabilize the activation loop would therefore increase GK activity. In contrast to the Gram negative structures that are available, the ATP binding site on Domain II does not go through an order-disorder transition when ATP binds. In the absence of ATP, this area is still stable, suggesting that directed targeting of the ATP binding site may be another way to differentially affect GK activity in Gram positive and Gram negative bacteria.

Interactions between the En. casseliflavus GK and En. faecalis HPr show that the presence of HPr in the crystallization setups affects the activation loop position. Although the phosphorylated histidine is necessary for increased activity in En. casseliflavus GK, the presence of HPr stabilizes the activation loop against Domain I. The stabilization of the activation loop is also affected by the stacking interactions that occur between the aromatic residues on the activation loop and aromatic residues on Domain I. Phosphorylation of His232 would create interactions between the activation loop and Domain I that would affect the OX dimer interfact and increase activity by closing the active cleft. HPr appears to not only phosphorylate His232 but also reposition the activation loop so that it is able to create the necessary contacts with Domain I. Unfortunately, the interaction is not visible in the current crystal structures, but further experiments using dynamic techniques such as SPR or NMR of the wild type and site-directed mutants may be able to further identify the regulation and interactions that occur between HPr and GK. Mutations around the activation loop that create different charges would be helpful to understand how phosphorylation of His232 modifies the interaction with HPr. A

phosphomimetic mutant of HPr may be necessary to create a stable interaction between HPr and GK, as HPr is phosphorylated when it activates GK. Our initial studies show that phosphorylation is not necessary for GK-HPr interaction in SPR experiments, but it may be necessary to have a negative charge to stabilize the interaction in crystallization conditions.

The *P. aeruginosa* GK adopts similar conformations to the *E. coli* structure in the glycerol-bound form. The active site residues are very similar between the *P. aeruginosa*, *E. coli*, and *En. casseliflavus* GK. As in the *En. casseliflavus* structures, the activation cleft closes more when bound to glycerol versus ethylene glycol. However, the ATP binding site on Domain II is disordered in both the glycerol-bound and ethylene glycol-bound *P. aeruginosa* structures. A structure of glycerol-bound P. aeruginosa GK with ATP analogues would show whether or not the ATP binding site goes through the same disorder-to-order transition that is seen in the E. coli GK structures when ATP or ADP is bound. If Domain II changes during ATP binding, this regulation step could be used to differentially inhibit *P. aeruginosa* GK activity and not affect the activity of *E. coli* GK. Inhibiting *P. aeruginosa* GK without affecting *E. coli* GK would allow targeting of the pathogen and preventing biofilm growth without affecting the growth of native *E. coli* strains in the human gut. Allowing the native *E. coli* colonies to remain would allow the normal balance of the gut to remain and prevent opportunistic pathogens from colonizing the gut. The tetrameric structure of the *P. aeruginosa* GK created by symmetry-related monomers is very close to the tetrameric structures of inactive *E. coli* GK, suggesting that these *P. aeruginosa* GK structures are indicative of inactive GK tetramers. Further research would be necessary to disrupt the OX dimer interface and determine the structures of the active *P. aeruginosa* GK.

The octameric structure of the hemoglobin mutants are affected by the presence of ligand in the molecule. When a diatomic ligand is present, there is less motion across the disulfide bridges than when no ligand or oxidization is present. This will affect the ability of the octameric hemoglobin to adopt specific conformations. However, this motion can be altered by changing the position of the mutations. For example, the αAsn78Cys mutant does not show the same type of variability in the disulfide bridges, and therefore it does not have as much variability in the tetramers as the βGly83Cys mutants. Restriction of the αAsn78Cys disulfide bridges is caused in part by the steric clashes that would result between the two tetramers if the disulfide bridges moved. The restriction on the disulfide bridges also means that the αAsn78Cys mutant would have to transition between the unliganded T state and the liganded R state in a more coordinated manner than the βGly83Cys mutants. The structures and crystallographic results suggest that the ligand stabilizes the βGly83Cys mutant relative tetramer motion.

The octameric hemoglobin adopts a tetramer conformation much like the structure crystallized in PEG, 1BBB [17], regardless of whether a hydroxyl or cyanide ligand is present in the binding pocket. The ability of the hydroxyl-bound mutants to be in the R state while still only partially liganded suggests the transition from the unliganded T state to the liganded R state is coordinated between the two linked tetramers, even though the cooperativity is similar to the cooperativity seen in adult human hemoglobin (HbA) [34, 128]. The lack of unliganded T state models and the presence of the oxidized form of the octameric hemoglobin limits our ability to give a broad analysis of the effects of the disulfide bridges on the T state and the transition between the unliganded and liganded form. The propensity of the mutants to form a structure with more rotation between the dimers when bound to ligand suggests that part of the higher oxygen affinity seen in the βGly83Cys is due in part to the protein being primed to form the R2

state by the disulfide bridge. A concerning point is that all of these structures are oxidized, including the liganded form. Determining the amount of oxidized octameric hemoglobin via spectra is necessary to determine the long-term potential of octameric hemoglobin as blood substitutes. Animal models using the octameric hemoglobin as a blood substitute have thus far shown no adverse effects and are successful at delivering oxygen (Dr. Chien Ho, personal communication), so it is possible that the reducing agents in plasma are sufficient to keep the octameric hemoglobin in a reduced state.

The Rap1B models demonstrate the portions of regulation that are determined by the phosphorylation of Ser179 rather than the presence of GTP. It has been postulated that the presence of GTP is the overarching regulator of Rap activity, but the Gly12Val-Ser179Asp-Cys181Gly and Gly12Val-Ser179Ala-Cys181Gly structures demonstrate a second level of regulation caused by the presence of a negative charge on residue 179. The negatively-charged Ser179Asp residue changes the interaction between the C-terminal loop and the core of the protein, presenting a binding area on the C-terminal loop that is inaccessible in the Ser179Ala mutant. Changes in the switch II region, which changes when GTP hydrolyzes in other Ras superfamilies and regulates protein-protein interactions, also show that the changes in the C-terminal charge can be transmitted through the protein. The changes suggest that the regulation caused by Ser179 phosphorylation is at least somewhat independent of the GTP-GDP status at the active sire. Further research should examine how the isoprenylation affects the activation of Rap1B, and examine the changes that occur with GTP binding, possibly with the use of different GTP analogues other than GTPγS.

All of these models demonstrate that the regulation of respiration is multifaceted across many biochemical pathways in ways that are just starting to identify. In many of the structural

models examined in this dissertation, the large changes were stabilized by charged interactions or hydrogen bonds. The binding of glycerol in the active site of GK creates hydrogen bonds between the glycerol and GK that close the active site and prime the GK for activation. Further interactions stabilize the activation loop that is 30 Å from the active site. The loop stabilization aligns the hydrophobic residues and further closes the active site for activation. Even in well-studied human hemoglobin, there are interactions between the heme pocket and the disulfide bridges introduced in the octameric hemoglobin that affect the overall ability to bind ligand.

Each of these components feeds into the respiratory system and affects the other components. Without hemoglobin, aerobic respiration and superoxide production would not occur because not enough oxygen would be present. Rap1B helps to regulate the immunological system and creates superoxide from the oxygen delivered by hemoglobin. The glycerol metabolism pathway supplies electrons to the electron transport chain that will end up utilizing oxygen to create water.

BIBLIOGRAPHY

1. Funke, O., *Uber das Milzveneneblut.* Zeitschrift fur rat. Medizin, 1851. **1**: p. 172-218.
2. Park, S.Y., et al., *1.25 A resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms.* J Mol Biol, 2006. **360**(3): p. 690-701.
3. Baldwin, J. and C. Chothia, *Haemoglobin: the structural changes related to ligand binding and its allosteric mechanism.* J Mol Biol, 1979. **129**(2): p. 175-220.
4. Baldwin, J.M., *The structure of human carbonmonoxy haemoglobin at 2.7 A resolution.* J Mol Biol, 1980. **136**(2): p. 103-28.
5. Muirhead, H. and M.F. Perutz, *Structure of Haemoglobin. A Three-Dimensional Fourier Synthesis of Reduced Human Haemoglobin at 5-5 a Resolution.* Nature, 1963. **199**: p. 633-8.
6. Safo, M.K. and D.J. Abraham, *The enigma of the liganded hemoglobin end state: a novel quaternary structure of human carbonmonoxy hemoglobin.* Biochemistry, 2005. **44**(23): p. 8347-59.
7. Shaanan, B., *Structure of human oxyhaemoglobin at 2.1 A resolution.* J Mol Biol, 1983. **171**(1): p. 31-59.
8. Kendrew, J.C., et al., *Structure of myoglobin: A three-dimensional Fourier synthesis at 2 A. resolution.* Nature, 1960. **185**(4711): p. 422-7.
9. Rossi-Fanelli, A. and E. Antonini, *Studies on the oxygen and carbon monoxide equilibria of human myoglobin.* Arch Biochem Biophys, 1958. **77**(2): p. 478-92.
10. Paoli, M., et al., *Crystal structure of T state haemoglobin with oxygen bound at all four haems.* J Mol Biol, 1996. **256**(4): p. 775-92.
11. Richard, V., G.G. Dodson, and Y. Mauguen, *Human deoxyhaemoglobin-2,3-diphosphoglycerate complex low-salt structure at 2.5 A resolution.* J Mol Biol, 1993. **233**(2): p. 270-4.
12. Tame, J.R. and B. Vallone, *The structures of deoxy human haemoglobin and the mutant Hb Tyralpha42His at 120 K.* Acta crystallographica. Section D, Biological crystallography, 2000. **56**(Pt 7): p. 805-11.
13. Joels, N. and L.G. Pugh, *The carbon monoxide dissociation curve of human blood.* J Physiol, 1958. **142**(1): p. 63-77.
14. Carlsen, E. and J.H. Comroe, Jr., *The rate of uptake of carbon monoxide and of nitric oxide by normal human erythrocytes and experimentally produced spherocytes.* J Gen Physiol, 1958. **42**(1): p. 83-107.
15. Lavalette, D., et al., *Linear free-energy relationships in binding of oxygen and carbon monoxide with heme model compounds and heme proteins.* Eur J Biochem, 1984. **145**(3): p. 555-65.

16. Kanaori, K., et al., *T-quaternary structure of oxy human adult hemoglobin in the presence of two allosteric effectors, L35 and IHP.* Biochim Biophys Acta, 2011. **1807**(10): p. 1253-61.
17. Silva, M.M., P.H. Rogers, and A. Arnone, *A third quaternary structure of human hemoglobin A at 1.7-A resolution.* J Biol Chem, 1992. **267**(24): p. 17248-56.
18. Ji, X., et al., *Alpha-subunit oxidation in T-state crystals of a sebacyl cross-linked human hemoglobin with unusual autoxidation properties.* Biophys Chem, 1998. **70**(1): p. 21-34.
19. Arnone, A., *X-ray diffraction study of binding of 2,3-diphosphoglycerate to human deoxyhaemoglobin.* Nature, 1972. **237**(5351): p. 146-9.
20. Arnone, A. and M.F. Perutz, *Structure of inositol hexaphosphate--human deoxyhaemoglobin complex.* Nature, 1974. **249**(452): p. 34-6.
21. Winslow, R.M., *The role of hemoglobin oxygen affinity in oxygen transport at high altitude.* Respir Physiol Neurobiol, 2007. **158**(2-3): p. 121-7.
22. Benesch, R. and R.E. Benesch, *Intracellular organic phosphates as regulators of oxygen release by haemoglobin.* Nature, 1969. **221**(5181): p. 618-22.
23. Chan, N.L., et al., *Crystallographic analysis of the interaction of nitric oxide with quaternary-T human hemoglobin.* Biochemistry, 2004. **43**(1): p. 118-32.
24. Perutz, M.F., *Regulation of oxygen affinity of hemoglobin: influence of structure of the globin on the heme iron.* Annu Rev Biochem, 1979. **48**: p. 327-86.
25. Yi, J., M.K. Safo, and G.B. Richter-Addo, *The nitrite anion binds to human hemoglobin via the uncommon O-nitrito mode.* Biochemistry, 2008. **47**(32): p. 8247-9.
26. Koshland, D.E., Jr., G. Nemethy, and D. Filmer, *Comparison of experimental binding data and theoretical models in proteins containing subunits.* Biochemistry, 1966. **5**(1): p. 365-85.
27. Monod, J., J. Wyman, and J.P. Changeux, *On the Nature of Allosteric Transitions: A Plausible Model.* J Mol Biol, 1965. **12**: p. 88-118.
28. Ackers, G.K. and J.M. Holt, *Asymmetric cooperativity in a symmetric tetramer: human hemoglobin.* J Biol Chem, 2006. **281**(17): p. 11441-3.
29. Holt, J.M. and G.K. Ackers, *Asymmetric distribution of cooperativity in the binding cascade of normal human hemoglobin. 2. Stepwise cooperative free energy.* Biochemistry, 2005. **44**(36): p. 11939-49.
30. Yonetani, T. and M. Laberge, *Protein dynamics explain the allosteric behaviors of hemoglobin.* Biochim Biophys Acta, 2008. **1784**(9): p. 1146-58.
31. Eaton, W.A., et al., *Evolution of allosteric models for hemoglobin.* IUBMB Life, 2007. **59**(8-9): p. 586-99.
32. Ackers, G.K., et al., *Molecular code for cooperativity in hemoglobin.* Science, 1992. **255**(5040): p. 54-63.
33. Auker, C.R. and R.M. McCarron, *US Navy experience with research on, and development of, hemoglobin-based oxygen carriers.* J Trauma, 2011. **70**(5 Suppl): p. S40-1.
34. Vasseur-Godbillon, C., et al., *Recombinant hemoglobin betaG83C-F41Y.* FEBS J, 2006. **273**(1): p. 230-41.
35. King, T.E. and S. Takemori, *Reconstitution of Respiratory Chain Enzyme Systems. Xiv. Reconstitution of Succinate-Cytochrome C Reductase from Soluble Succinate Dehydrogenase and the Cytochrome B-C1 Particle.* J Biol Chem, 1964. **239**: p. 3559-69.

36. Saier, M.H., Jr. and S. Roseman, *Sugar transport. 2nducer exclusion and regulation of the melibiose, maltose, glycerol, and lactose transport systems by the phosphoenolpyruvate:sugar phosphotransferase system.* J Biol Chem, 1976. **251**(21): p. 6606-15.
37. Deutscher, J., C. Francke, and P.W. Postma, *How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria.* Microbiol Mol Biol Rev, 2006. **70**(4): p. 939-1031.
38. Bizzini, A., et al., *Glycerol is metabolized in a complex and strain-dependent manner in Enterococcus faecalis.* J Bacteriol, 2010. **192**(3): p. 779-85.
39. Halbedel, S., C. Hames, and J. Stulke, *In vivo activity of enzymatic and regulatory components of the phosphoenolpyruvate:sugar phosphotransferase system in Mycoplasma pneumoniae.* J Bacteriol, 2004. **186**(23): p. 7936-43.
40. Joseph, B., et al., *Identification of Listeria monocytogenes genes contributing to intracellular replication by expression profiling and mutant screening.* J Bacteriol, 2006. **188**(2): p. 556-68.
41. Pilo, P., et al., *A metabolic enzyme as a primary virulence factor of Mycoplasma mycoides subsp. mycoides small colony.* J Bacteriol, 2005. **187**(19): p. 6824-31.
42. Schweizer, H. and T.J. Larson, *Cloning and characterization of the aerobic sn-glycerol-3-phosphate dehydrogenase structural gene glpD of Escherichia coli K-12.* J Bacteriol, 1987. **169**(2): p. 507-13.
43. Walz, A.C., et al., *Aerobic sn-glycerol-3-phosphate dehydrogenase from Escherichia coli binds to the cytoplasmic membrane through an amphipathic alpha-helix.* Biochem J, 2002. **365**(Pt 2): p. 471-9.
44. Colussi, T., et al., *Structure of alpha-glycerophosphate oxidase from Streptococcus sp.: a template for the mitochondrial alpha-glycerophosphate dehydrogenase.* Biochemistry, 2008. **47**(3): p. 965-77.
45. Gay, P., et al., *Carbohydrate metabolism and transport in Bacillus subtilis. A study of ctr mutations.* Mol Gen Genet, 1973. **121**(4): p. 355-68.
46. Romano, A.H., et al., *Physiological studies on regulation of glycerol utilization by the phosphoenolpyruvate:sugar phosphotransferase system in Enterococcus faecalis.* J Bacteriol, 1990. **172**(12): p. 6741-8.
47. Beijer, L., et al., *The glpP and glpF genes of the glycerol regulon in Bacillus subtilis.* J Gen Microbiol, 1993. **139**(2): p. 349-59.
48. Darbon, E., et al., *Glycerol transport and phosphoenolpyruvate-dependent enzyme I- and HPr-catalysed phosphorylation of glycerol kinase in Thermus flavus.* Microbiology, 1999. **145 (Pt 11)**: p. 3205-12.
49. Spoering, A.L., M. Vulic, and K. Lewis, *GlpD and PlsB participate in persister cell formation in Escherichia coli.* J Bacteriol, 2006. **188**(14): p. 5136-44.
50. Weissenborn, D.L., N. Wittekindt, and T.J. Larson, *Structure and regulation of the glpFK operon encoding glycerol diffusion facilitator and glycerol kinase of Escherichia coli K-12.* J Biol Chem, 1992. **267**(9): p. 6122-31.
51. Fu, D., et al., *Structure of a glycerol-conducting channel and the basis for its selectivity.* Science, 2000. **290**(5491): p. 481-6.
52. Lin, E.C., *Glycerol dissimilation and its regulation in bacteria.* Annu Rev Microbiol, 1976. **30**: p. 535-78.

53. Ormo, M., C.E. Bystrom, and S.J. Remington, *Crystal structure of a complex of Escherichia coli glycerol kinase and an allosteric effector fructose 1,6-bisphosphate.* Biochemistry, 1998. **37**(47): p. 16565-72.
54. Hurley, J.H., et al., *Structure of the regulatory complex of Escherichia coli IIIGlc with glycerol kinase.* Science, 1993. **259**(5095): p. 673-7.
55. Feese, M., et al., *Cation-promoted association of a regulatory and target protein is controlled by protein phosphorylation.* Proc Natl Acad Sci U S A, 1994. **91**(9): p. 3544-8.
56. Postma, P.W., et al., *Interaction between IIIGlc of the phosphoenolpyruvate:sugar phosphotransferase system and glycerol kinase of Salmonella typhimurium.* J Bacteriol, 1984. **158**(1): p. 351-3.
57. Bystrom, C.E., et al., *Crystal structures of Escherichia coli glycerol kinase variant S58-->W in complex with nonhydrolyzable ATP analogues reveal a putative active conformation of the enzyme as a result of domain motion.* Biochemistry, 1999. **38**(12): p. 3508-18.
58. Liu, W.Z., et al., *Escherichia coli glycerol kinase: role of a tetramer interface in regulation by fructose 1,6-bisphosphate and phosphotransferase system regulatory protein IIIglc.* Biochemistry, 1994. **33**(33): p. 10120-6.
59. Charrier, V., et al., *Cloning and sequencing of two enterococcal glpK genes and regulation of the encoded glycerol kinases by phosphoenolpyruvate-dependent, phosphotransferase system-catalyzed phosphorylation of a single histidyl residue.* J Biol Chem, 1997. **272**(22): p. 14166-74.
60. Shima, F., et al., *Structural basis for conformational dynamics of GTP-bound Ras protein.* J Biol Chem, 2010. **285**(29): p. 22696-705.
61. Tong, L.A., et al., *Crystal structures at 2.2 A resolution of the catalytic domains of normal ras protein and an oncogenic mutant complexed with GDP.* J Mol Biol, 1991. **217**(3): p. 503-16.
62. Hall, B.E., D. Bar-Sagi, and N. Nassar, *The structural basis for the transition from Ras-GTP to Ras-GDP.* Proc Natl Acad Sci U S A, 2002. **99**(19): p. 12138-42.
63. Fehrenbacher, N., D. Bar-Sagi, and M. Philips, *Ras/MAPK signaling from endomembranes.* Mol Oncol, 2009. **3**(4): p. 297-307.
64. Whelan, J.T., et al., *Post-transcriptional regulation of the Ras-ERK/MAPK signaling pathway.* J Cell Physiol, 2012. **227**(3): p. 1235-41.
65. Cerione, R.A., *Cdc42: new roads to travel.* Trends Cell Biol, 2004. **14**(3): p. 127-32.
66. Johnson, D.I., *Cdc42: An essential Rho-type GTPase controlling eukaryotic cell polarity.* Microbiol Mol Biol Rev, 1999. **63**(1): p. 54-105.
67. Nobes, C.D. and A. Hall, *Rho, rac and cdc42 GTPases: regulators of actin structures, cell adhesion and motility.* Biochem Soc Trans, 1995. **23**(3): p. 456-9.
68. Der, C.J., B.T. Pan, and G.M. Cooper, *rasH mutants deficient in GTP binding.* Mol Cell Biol, 1986. **6**(9): p. 3291-4.
69. Hu, Y., et al., *K-ras(G12V) transformation leads to mitochondrial dysfunction and a metabolic switch from oxidative phosphorylation to glycolysis.* Cell Res, 2012. **22**(2): p. 399-412.
70. Babior, B.M., *The respiratory burst oxidase.* Adv Enzymol Relat Areas Mol Biol, 1992. **65**: p. 49-95.
71. Segal, A.W. and O.T. Jones, *Novel cytochrome b system in phagocytic vacuoles of human granulocytes.* Nature, 1978. **276**(5687): p. 515-7.

72. Maly, F.E., et al., *Activated or dominant inhibitory mutants of Rap1A decrease the oxidative burst of Epstein-Barr virus-transformed human B lymphocytes.* J Biol Chem, 1994. **269**(29): p. 18743-6.
73. Hoshijima, M., et al., *Phosphorylation by cyclic AMP-dependent protein kinase of a human platelet Mr 22,000 GTP-binding protein (smg p21) having the same putative effector domain as the ras gene products.* Biochem Biophys Res Commun, 1988. **157**(3): p. 851-60.
74. Kawata, M., et al., *Phosphorylation of smg p21, a ras p21-like GTP-binding protein, by cyclic AMP-dependent protein kinase in a cell-free system and in response to prostaglandin E1 in intact human platelets.* J Biol Chem, 1989. **264**(26): p. 15688-95.
75. Siess, W., D.A. Winegar, and E.G. Lapetina, *Rap1-B is phosphorylated by protein kinase A in intact human platelets.* Biochem Biophys Res Commun, 1990. **170**(2): p. 944-50.
76. Siess, W. and E.G. Lapetina, *Functional relationship between cyclic AMP-dependent protein phosphorylation and platelet inhibition.* Biochem J, 1990. **271**(3): p. 815-9.
77. Lapetina, E.G., et al., *A ras-related protein is phosphorylated and translocated by agonists that increase cAMP levels in human platelets.* Proc Natl Acad Sci U S A, 1989. **86**(9): p. 3131-4.
78. Lerosey, I., et al., *The cAMP-dependent protein kinase phosphorylates the rap1 protein in vitro as well as in intact fibroblasts, but not the closely related rap2 protein.* Biochem Biophys Res Commun, 1991. **175**(2): p. 430-6.
79. Quilliam, L.A., et al., *Rap1A is a substrate for cyclic AMP-dependent protein kinase in human neutrophils.* J Immunol, 1991. **147**(5): p. 1628-35.
80. Buss, J.E., et al., *The COOH-terminal domain of the Rap1A (Krev-1) protein is isoprenylated and supports transformation by an H-Ras:Rap1A chimeric protein.* Mol Cell Biol, 1991. **11**(3): p. 1523-30.
81. Maltese, W.A., et al., *Post-translational modification of low molecular mass GTP-binding proteins by isoprenoid.* J Biol Chem, 1990. **265**(4): p. 2148-55.
82. Ohtsuka, T., et al., *Activation of brain B-Raf protein kinase by Rap1B small GTP-binding protein.* J Biol Chem, 1996. **271**(3): p. 1258-61.
83. Hata, Y., et al., *Enhancement of the actions of smg p21 GDP/GTP exchange protein by the protein kinase A-catalyzed phosphorylation of smg p21.* J Biol Chem, 1991. **266**(10): p. 6571-7.
84. Lazarowski, E.R., J.C. Lacal, and E.G. Lapetina, *Agonist-induced phosphorylation of an immunologically ras-related protein in human erythroleukemia cells.* Biochem Biophys Res Commun, 1989. **161**(3): p. 972-8.
85. Minato, N., K. Kometani, and M. Hattori, *Regulation of immune responses and hematopoiesis by the Rap1 signal.* Adv Immunol, 2007. **93**: p. 229-64.
86. Ihara, K., et al., *Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue.* J Biol Chem, 1998. **273**(16): p. 9656-66.
87. Nassar, N., et al., *The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue.* Nature, 1995. **375**(6532): p. 554-60.
88. Rehmann, H., et al., *Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B.* Nature, 2008. **455**(7209): p. 124-7.
89. Scrima, A., et al., *The Rap-RapGAP complex: GTP hydrolysis without catalytic glutamine and arginine residues.* EMBO J, 2008. **27**(7): p. 1145-53.

90. Verrotti, A.C., et al., *RAS residues that are distant from the GDP binding site play a critical role in dissociation factor-stimulated release of GDP.* EMBO J, 1992. **11**(8): p. 2855-62.
91. Coulon, C., et al., *Chemical analysis of cellular and extracellular carbohydrates of a biofilm-forming strain Pseudomonas aeruginosa PA14.* PLoS One, 2010. **5**(12): p. e14220.
92. Frohlich, K.M., et al., *Rickettsia prowazekii uses an sn-glycerol-3-phosphate dehydrogenase and a novel dihydroxyacetone phosphate transport system to supply triose phosphate for phospholipid biosynthesis.* J Bacteriol, 2010. **192**(17): p. 4281-8.
93. Novotny, M.J., et al., *Allosteric regulation of glycerol kinase by enzyme IIIglc of the phosphotransferase system in Escherichia coli and Salmonella typhimurium.* J Bacteriol, 1985. **162**(2): p. 810-6.
94. Deutscher, J., B. Bauer, and H. Sauerwald, *Regulation of glycerol metabolism in Enterococcus faecalis by phosphoenolpyruvate-dependent phosphorylation of glycerol kinase catalyzed by enzyme I and HPr of the phosphotransferase system.* J Bacteriol, 1993. **175**(12): p. 3730-3.
95. Parsonage, D., et al., *The soluble alpha-glycerophosphate oxidase from Enterococcus casseliflavus. Sequence homology with the membrane-associated dehydrogenase and kinetic analysis of the recombinant enzyme.* J Biol Chem, 1998. **273**(37): p. 23812-22.
96. Austin, D. and T.J. Larson, *Nucleotide sequence of the glpD gene encoding aerobic sn-glycerol 3-phosphate dehydrogenase of Escherichia coli K-12.* J Bacteriol, 1991. **173**(1): p. 101-7.
97. Wehtje, C., et al., *Mutations in the glycerol kinase gene restore the ability of a ptsGHI mutant of Bacillus subtilis to grow on glycerol.* Microbiology, 1995. **141 (Pt 5)**: p. 1193-8.
98. Gasteiger, E., et al., *Protein Identification and Analysis Tools on the ExPASy Server*, in *The Proteomics Protocols Handbook*, J.M. Walker, Editor 2005, Humana Press. p. 571-607.
99. Yeh, J.I., U. Chinte, and S.C. Du, *Structure of glycerol-3-phosphate dehydrogenase, an essential monotopic membrane enzyme involved in respiration and metabolism.* Proc Natl Acad Sci U S A, 2008. **105**(9): p. 3280-3285.
100. Yeh, J.I., et al., *Structures of enterococcal glycerol kinase in the absence and presence of glycerol: correlation of conformation to substrate binding and a mechanism of activation by phosphorylation.* Biochemistry, 2004. **43**(2): p. 362-73.
101. Yeh, J.I., et al., *Structural characterizations of glycerol kinase: unraveling phosphorylation-induced long-range activation.* Biochemistry, 2009. **48**(2): p. 346-56.
102. Pflugrath, J.W., *The finer things in X-ray diffraction data collection.* Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 10): p. 1718-25.
103. Otwinowski, Z. and W. Minor, *Processing of X-ray Diffraction Data COllected in Oscillation Mode.* Methods Enzymol, 1997. **276**: p. 307-326.
104. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
105. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination.* Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 905-21.

106. *The CCP4 suite: programs for protein crystallography.* Acta Crystallogr D Biol Crystallogr, 1994. **50**(Pt 5): p. 760-3.
107. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
108. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
109. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
110. Richards, F.M. and C.E. Kundrot. *DDMP*. 1988; Available from: http://www.csb.yale.edu/userguides/datamanip/ddmp/ddmp_descrip.html.
111. Kleywegt, G.J., *Use of non-crystallographic symmetry in protein structure refinement.* Acta crystallographica. Section D, Biological crystallography, 1996. **52**(Pt 4): p. 842-57.
112. Kleywegt, G.J. and T.A. Jones, *Detecting folding motifs and similarities in protein structures.* Methods in enzymology, 1997. **277**: p. 525-45.
113. Feese, M.D., et al., *Glycerol kinase from Escherichia coli and an Ala65-->Thr mutant: the crystal structures reveal conformational changes with implications for allosteric regulation.* Structure, 1998. **6**(11): p. 1407-18.
114. Anderson, M.J., et al., *Crystal structure of a hyperactive Escherichia coli glycerol kinase mutant Gly230 --> Asp obtained using microfluidic crystallization devices.* Biochemistry, 2007. **46**(19): p. 5722-31.
115. Voss, N.R., et al., *The geometry of the ribosomal polypeptide exit tunnel.* J Mol Biol, 2006. **360**(4): p. 893-906.
116. Perutz, M.F., *Preparation of Haemoglobin Crystals.* Journal of Crystal Growth, 1968. **2**: p. 54-56.
117. Gong, Q., et al., *Quaternary structure of carbonmonoxyhemoglobins in solution: structural changes induced by the allosteric effector inositol hexaphosphate.* Biochemistry, 2006. **45**(16): p. 5140-8.
118. Lukin, J.A., et al., *Quaternary structure of hemoglobin in solution.* Proc Natl Acad Sci U S A, 2003. **100**(2): p. 517-20.
119. Fablet, C., et al., *Stable octameric structure of recombinant hemoglobin alpha(2)beta(2)83 Gly-->Cys.* Protein Sci, 2003. **12**(4): p. 690-5.
120. Fronticelli, C., et al., *Molecular engineering of a polymer of tetrameric hemoglobins.* Proteins, 2001. **44**(3): p. 212-22.
121. Jahr, J.S., A.S. Akha, and R.J. Holtby, *Crosslinked, Polymerized, and PEG-conjugated Hemoglobin-based Oxygen Carriers: Clinical Safety and Efficacy of Recent and Current Products.* Curr Drug Discov Technol, 2011.
122. Manjula, B.N., et al., *Site-specific PEGylation of hemoglobin at Cys-93(beta): correlation between the colligative properties of the PEGylated protein and the length of the conjugated PEG chain.* Bioconjug Chem, 2003. **14**(2): p. 464-72.
123. Rameez, S. and A.F. Palmer, *Simple method for preparing poly(ethylene glycol)-surface-conjugated liposome-encapsulated hemoglobins: physicochemical properties, long-term storage stability, and their reactions with O2, CO, and NO.* Langmuir, 2011. **27**(14): p. 8829-40.
124. Sakai, H., et al., *Poly(ethylene glycol)-conjugation and deoxygenation enable long-term preservation of hemoglobin-vesicles as oxygen carriers in a liquid state.* Bioconjug Chem, 2000. **11**(3): p. 425-32.

125. Vandegriff, K.D., et al., *Oxidation and haem loss kinetics of poly(ethylene glycol)-conjugated haemoglobin (MP4): dissociation between in vitro and in vivo oxidation rates.* Biochem J, 2006. **399**(3): p. 463-71.
126. Zhang, W., et al., *A novel hemoglobin-based oxygen carrier, polymerized porcine hemoglobin, inhibits H(2)O(2)-induced cytotoxicity of endothelial cells.* Artif Organs, 2012. **36**(2): p. 151-60.
127. Chauvierre, C., et al., *Artificial oxygen carrier based on polysaccharides-poly(alkylcyanoacrylates) nanoparticle templates.* Biomaterials, 2010. **31**(23): p. 6069-74.
128. Brillet, T., et al., *Interaction of haptoglobin with haemoglobin octamers based on the mutation alphaAsn78Cys or betaGly83Cys.* American Journal of Molecular Biology, 2012. **In press**.
129. Marden, M.C., et al., *Modulation of the oxygen affinity of cobalt-porphyrin by globin.* FEBS Lett, 2000. **472**(2-3): p. 221-4.
130. Olson, J.S., et al., *Protein engineering strategies for designing more stable hemoglobin-based blood substitutes.* Artif Cells Blood Substit Immobil Biotechnol, 1997. **25**(1-2): p. 227-41.
131. Bucci, E., et al., *Positive and negative cooperativities at subsequent steps of oxygenation regulate the allosteric behavior of multistate sebacylhemoglobin.* Biochemistry, 1996. **35**(11): p. 3418-25.
132. Schumacher, M.A., et al., *Allosteric intermediates indicate R2 is the liganded hemoglobin end state.* Proc Natl Acad Sci U S A, 1997. **94**(15): p. 7841-4.
133. Park, S.Y., et al., *Crystal structures of unliganded and half-liganded human hemoglobin derivatives cross-linked between Lys 82beta1 and Lys 82beta2.* Biochemistry, 2004. **43**(27): p. 8711-7.
134. Benesch, R.E. and S. Kwong, *The stability of the heme-globin linkage in some normal, mutant, and chemically modified hemoglobins.* J Biol Chem, 1990. **265**(25): p. 14881-5.
135. Walder, J.A., R.Y. Walder, and A. Arnone, *Development of antisickling compounds that chemically modify hemoglobin S specifically within the 2,3-diphosphoglycerate binding site.* J Mol Biol, 1980. **141**(2): p. 195-216.
136. Lui, F.E. and R. Kluger, *Enhancing nitrite reductase activity of modified hemoglobin: bis-tetramers and their PEGylated derivatives.* Biochemistry, 2009. **48**(50): p. 11912-9.
137. Arnone, A., R.E. Benesch, and R. Benesch, *Structure of human deoxyhemoglobin specifically modified with pyridoxal compounds.* J Mol Biol, 1977. **115**(4): p. 627-42.
138. Johnson, M.B., J.G. Adamson, and A.G. Mauk, *Functional comparison of specifically cross-linked hemoglobins biased toward the R and T states.* Biophys J, 1998. **75**(6): p. 3078-84.
139. Bucci, E., et al., *Effect of intramolecular cross-links on the enthalpy and quaternary structure of the intermediates of oxygenation of human hemoglobin.* Biochemistry, 1993. **32**(14): p. 3519-26.
140. Vandegriff, K.D., et al., *Carbon dioxide binding to human hemoglobin cross-linked between the alpha chains.* J Biol Chem, 1991. **266**(5): p. 2697-700.
141. Vandegriff, K.D., et al., *Determination of the rate and equilibrium constants for oxygen and carbon monoxide binding to R-state human hemoglobin cross-linked between the alpha subunits at lysine 99.* J Biol Chem, 1991. **266**(26): p. 17049-59.

142. Jones, R.T., et al., *A doubly cross-linked human hemoglobin. Effects of cross-links between different subunits.* J Biol Chem, 1996. **271**(2): p. 675-80.
143. Chatterjee, R., et al., *Isolation and characterization of a new hemoglobin derivative cross-linked between the alpha chains (lysine 99 alpha 1----lysine 99 alpha 2).* J Biol Chem, 1986. **261**(21): p. 9929-37.
144. Marden, M.C., et al., *Control of the allosteric equilibrium of hemoglobin by cross-linking agents.* Protein Sci, 2002. **11**(6): p. 1376-83.
145. Hu, D. and R. Kluger, *Functional cross-linked hemoglobin bis-tetramers: geometry and cooperativity.* Biochemistry, 2008. **47**(47): p. 12551-61.
146. Nagel, R.L. and Q.H. Gibson, *The binding of hemoglobin to haptoglobin and its relation to subunit dissociation of hemoglobin.* J Biol Chem, 1971. **246**(1): p. 69-73.
147. Blackwell, R.Q., C.S. Liu, and C.L. Wang, *Hemoglobin Ta-Li: 83 Gly leads to Cys.* Biochim Biophys Acta, 1971. **243**(3): p. 467-74.
148. Rai, D.K., et al., *Characterization of the elusive disulfide bridge forming human Hb variant: Hb Ta-Li beta83 (EF7)Gly --> Cys by electrospray mass spectrometry.* J Am Soc Mass Spectrom, 2002. **13**(2): p. 187-91.
149. Patrinos, G.P., et al., *Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies.* Nucleic Acids Res, 2004. **32**(32): p. D537-541.
150. Hill, A.V., *The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves.* J Physiol, 1910. **40**: p. iv-vii.
151. Shen, T.J., et al., *Production of unmodified human adult hemoglobin in Escherichia coli.* Proc Natl Acad Sci U S A, 1993. **90**(17): p. 8108-12.
152. McCoy, A.J., et al., *Likelihood-enhanced fast translation functions.* Acta Crystallogr D Biol Crystallogr, 2005. **61**(Pt 4): p. 458-64.
153. Winn, M.D., G.N. Murshudov, and M.Z. Papiz, *Macromolecular TLS refinement in REFMAC at moderate resolutions.* Methods Enzymol, 2003. **374**: p. 300-21.
154. Adams, P.D., et al., *PHENIX: building new software for automated crystallographic structure determination.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 11): p. 1948-54.
155. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W375-83.
156. Brunger, A.T., *Free R value: a novel statistical quantity for assessing the accuracy of crystal structures.* Nature, 1992. **355**(6359): p. 472-5.
157. Vaguine, A.A., J. Richelle, and S.J. Wodak, *SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model.* Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 1): p. 191-205.
158. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* J. App. Cryst., 1993. **26**: p. 283-291.
159. Park, S.Y., et al., *1.25 A resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms.* Journal of molecular biology, 2006. **360**(3): p. 690-701.
160. Silva, M.M., P.H. Rogers, and A. Arnone, *A third quaternary structure of human hemoglobin A at 1.7-A resolution.* The Journal of biological chemistry, 1992. **267**(24): p. 17248-56.
161. Paoli, M., et al., *Crystal structure of T state haemoglobin with oxygen bound at all four haems.* Journal of molecular biology, 1996. **256**(4): p. 775-92.

162. Bos, J.L., J. de Rooij, and K.A. Reedquist, *Rap1 signalling: adhering to new models.* Nat Rev Mol Cell Biol, 2001. **2**(5): p. 369-77.
163. Adari, H., et al., *Guanosine triphosphatase activating protein (GAP) interacts with the p21 ras effector binding domain.* Science, 1988. **240**(4851): p. 518-21.
164. Kitayama, H., et al., *Genetic analysis of the Kirsten-ras-revertant 1 gene: potentiation of its tumor suppressor activity by specific point mutations.* Proc Natl Acad Sci U S A, 1990. **87**(11): p. 4284-8.
165. Boettner, B. and L. Van Aelst, *Control of cell adhesion dynamics by Rap1 signaling.* Curr Opin Cell Biol, 2009. **21**(5): p. 684-93.